Dynamic evolution of small signaling peptide compensation in plant stem cell control

Choon-Tak Kwon[1,2], Lingli Tang[3,4], Xingang Wang[1], Iacopo Gentile[1], Anat Hendelman[1], Gina Robitaille[1,7], Joyce Van Eck[5,6], Cao Xu[3,4*], Zachary B. Lippman[1,7,*]

[1] School of Biological Sciences, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA.

[2] Department of Horticultural Biotechnology, Kyung Hee University, Yongin 17104, Republic of Korea

[3] State Key Laboratory of Plant Genomics, National Center for Plant Gene Research, CAS-JIC Centre of Excellence for Plant and Microbial Science (CEPAMS), Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, China.

[4] University of Chinese Academy of Sciences, Beijing, China.

[5] Boyce Thompson Institute, Ithaca, NY 14853, USA

[6] Plant Breeding and Genetics Section, School of Integrative Plant Science, Cornell University, Ithaca, NY 14853, USA

[7] Howard Hughes Medical Institute, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA

*e-mail: caoxu@genetics.ac.cn, lippman@cshl.edu

Keywords: Paralogs, Redundancy, Epistasis, Signaling Peptide, *cis*-regulatory, Meristem, Tomato, Solanaceae, CRISPR

Gene duplications are a hallmark of plant genome evolution and a foundation for genetic interactions that shape phenotypic diversity[1–5]. Compensation is a major form of paralog interaction[6–8], but how compensation relationships change as allelic variation accumulates is unknown. Here, we leveraged genomics and genome editing across the Solanaceae family to capture the evolution of compensating paralogs. Mutations in the stem cell regulator *CLV3* cause floral organs to overproliferate in many plants[9–11]. In tomato, this phenotype is partially suppressed by transcriptional upregulation of a closely related paralog[12]. Tobacco lost this paralog, resulting in no compensation and extreme *clv3* phenotypes. Strikingly, the paralogs of petunia and groundcherry nearly completely suppress *clv3*, indicating a potent ancestral state of compensation. Cross-species transgenic complementation analyses show this potent compensation partially degenerated in tomato due to a single amino acid change in the paralog and *cis*-regulatory variation that limits its transcriptional upregulation. Our findings show how genetic interactions are remodeled following duplications, and suggest that dynamic paralog evolution is widespread over short time scales and impacts phenotypic variation from natural and engineered mutations.

Gene duplications arise from whole genome and small-scale duplications and are pervasive in plant genomes[3,5,13,14]. Paralogs that emerge from duplications are completely redundant, which allows genetic variation to accumulate under relaxed selection[3,5]. This mutational drift can diversify paralog relationships through gene loss (pseudogenization), partitioning of ancestral functions (subfunctionalization), or gain of novel functions (neofunctionalization)[1,3,5,15]. Another prominent but less understood path of paralog evolution leads to "active compensation", a form of redundancy where one or more paralogs are transcriptionally upregulated to substitute for the compromised activity of another[6,16,17]. Such relationships provide robustness against genetic or environmental change and may be under selection[18,19]. However, an often underappreciated paradox is that while duplications initially provide redundancy, they also promote new genetic variation through relaxed purifying selection[18,20,21] . Such variation, which can accumulate across both coding and *cis*-regulatory sequences, is the foundation for the broadly studied end-points of paralog diversification. What remains unclear is how such diversification modifies paralog functional relationships as species diversify over shorter time frames. This is because functional dissections of paralogs have been limited to within individual systems or between a few widely

divergent species, and thus have failed to capture the trajectories and functional consequences of evolving compensatory relationships following lineage-specific ancestral duplications[6,12,14].

CLAVATA3/EMBRYO-SURROUNDING REGION-RELATED (*CLE*) genes comprise an important gene family in plants encoding small-signaling peptides with diverse roles in growth and development[22,23]. CLE peptides are 12- or 13-residue glycopeptides processed from pre-propeptides[23,24]. The number and functional relationships, including redundancy, of CLE family members, vary considerably between distantly related species, due to lineage-specific duplications and variation in paralog retention and diversification[22]. However, the founding member from *Arabidopsis thaliana* (arabidopsis), CLAVATA3 (CLV3), is deeply conserved[9,25]. The CLV3 dodecapeptide is a ligand for the leucine-rich receptor kinase CLV1 and related receptors, and functions in a negative feedback circuit with WUSCHEL (WUS), a homeobox transcription factor that promotes stem cell production in shoot meristems[10,11]. Mutations in *CLV3* and its orthologs in many species cause meristem enlargement, which leads to tissue and organ overproliferation, or fasciation, phenotypes, especially in flowers[9,10]. We previously showed that *clv3* mutations in the divergent species arabidopsis, *Zea mays* (maize), and *Solanum lycopersicum* (tomato) are buffered through redundancy, but through different mechanisms[12]. In arabidopsis, multiple *CLE* family members partially suppress *clv3* without changing their expression[12]. In contrast to this "passive compensation", a similar partial suppression of *clv3* mutations in maize (*zmcle7*) and tomato (*slclv3*) is achieved by active compensation from closely related *CLV3* paralogs[12]. Though the mechanism of compensation is shared between maize and tomato, the paralogs involved arose through lineage-specific duplications, indicating independent evolution of active compensation. Thus, it remains unclear how states of active compensation are achieved in any lineage and whether they remain stable or continue to evolve as species diversify.

With several genetically tractable species, closely related Solanaceae family members comprise a useful system to track the evolution of the compensation relationship between *CLV3* and its paralog. The compensating paralog in tomato, *SlCLE9*, originated from a duplication event just prior to diversification of the Solanales[12]. CRISPR-Cas9 engineered *slcle9* mutations result in normal plants, but strongly enhance *slclv3* due to loss of active compensation (**Fig. 1a-c**). Interestingly, our synteny analysis of 29 Solanaceae genomes capturing ~30 million years of evolution revealed several species that partially or completely lost their *SlCLE9* orthologs (**Fig. 1d and Supplementary Table 1**)[12]. For example, whereas *Physalis grisea* (groundcherry) and

Petunia hybrida (petunia) have *SlCLE9* orthologs, *Capsicum annuum* (pepper) harbors only fragments of an *SlCLE9* ortholog, indicating pseudogenization (**Fig. 1d and Supplementary Table 1**)[12]. Both *S. tuberosum* (potato) and *S. melongena* (eggplant) lack *SlCLE9* orthologs entirely, and this presence-absence variation extends to the genus level; in *Nicotiana* (tobacco), the *SlCLE9* orthologs in *N. tabacum* and *N. benthamiana* were retained or pseudogenized, respectively (**Fig. 1d and Supplementary Table 1**).

Since active compensation is typically mediated by the existence of a close paralog[6,16], we predicted that species that lost their *SlCLE9* orthologs would lack active compensation. However, in such species, compensation could also have evolved from one or more *CLE* homologs, which could potentially compensate passively (i.e. without transcriptional upregulation), as found in the Brassicaceae species *Arabidopsis thaliana*[12]. We tested compensation in the allotetraploid *N. benthamiana*, where CRISPR-Cas9 genome editing is highly efficient, but brings an added layer of genetic complexity from having two sub-genome copies (homeologs) of all genes, including *NbCLV3* (*NbCLV3a* and *NbCLV3b*)[26]. To test for loss of compensation in this species, we designed a multiplex CRISPR-Cas9 construct with eight guide RNAs (gRNAs) designed to target *NbCLV3a* and *NbCLV3b* (four gRNAs each; **Fig 1e**). We obtained five first-generation transgenic (T_0) plants, and unsurprisingly, all were chimeric (**Extended Data Fig. 1a-c**). Three of these plants exhibited severe fasciation phenotypes like tomato *slclv3 slcle9* double mutants, including thick stems and extreme overproliferation of floral organs, whereas the other two plants were less fasciated (**Extended Data Fig. 1c-d**). Though all plants were chimeric for mutations in *NbCLV3a* and *NbCLV3b*, sequencing showed the three strongest mutants carried only mutated alleles of both genes, suggesting a null-equivalent phenotype similar to tomato *slclv3 slcle9* double mutants (**Fig. 1c and Extended Data Fig. 1a-c**). Though the severity of the floral fasciation in the strongest T_0 plants precluded recovery of mutant seeds, these observations supported the absence of active compensation in *N. benthamiana.* Importantly, we further validated these results in T_1 segregating lines derived from the weaker T_0 plants, which fortuitously provided progeny populations that carried null alleles of *nbclv3b* and segregated for a null allele of *nbclv3a* (**Fig. 1e-i**). We used these populations to isolate *nbclv3a/b* allotetraploid mutants and showed that meristems were more than twice as large in these plants compared to *nbclv3b* single mutants and wild-type controls (**Fig. 1j, k**). Together, these results show that active compensation in the regulation of meristem

maintenance was lost in *N. benthamiana* and also supports that conservation of active compensation in the Solanaceae requires retention of *SlCLE9* orthologs.

We next asked if compensation varies in lineages that retained their *SlCLE9* orthologs, and where allelic variation between these lineages could affect paralog function. Orthologous CLE pre-propeptide sequences are highly variable between species, but their dodecapeptides are more conserved[22,23]. Indeed, while *SlCLV3* and *SlCLE9* ortholog dodecapeptide sequences were nearly invariant in the Solanaceae, we found widespread variation in the coding and putative cis-regulatory regions of both genes, the latter determined by conserved non-coding sequence (CNS) analyses (**Extended Data Fig. 2 and Supplementary Table 1**). To assess active compensation in other Solanaceae species carrying *SlCLE9* orthologs, we took advantage of established CRISPR-Cas9 genome editing in petunia (**Fig. 2a**). Strikingly, the phenotypes of independently derived *phclv3* null mutants were both substantially weaker than tomato *slclv3* mutants (**Fig. 1b, 2b-d**). Although the primary shoot meristem was larger than wild-type meristems, 80% of *phclv3* flowers produced wild-type organ numbers (**Fig. 2c-f**). Given that multiple attempts to generate *pgcle9* mutants were unsuccessful, we micro-dissected *phclv3* meristems for RNA-sequencing to profile differentially expressed genes due to mutation of *PhCLV3*. Notably, out of all petunia *CLE* family members only *PhCLE9* was dramatically upregulated (>15-fold) (**Fig. 2g, h and Supplementary Table 2**), consistent with *SlCLE9* upregulation in tomato *slclv3* mutants and suggesting active compensation in petunia is mediated by *PhCLE9* and is stronger than in tomato.

Conservation of CLE dodecapeptide sequences is critical for proper ligand folding and receptor binding[27,28]. A single amino acid at position 6 distinguishes the petunia PhCLE9 and tomato SlCLE9 dodecapeptides, and a deeper analysis of conservation revealed that all species from tomato and its wild relatives through *Jaltomata sinuosa* have a serine at this position, whereas all other Solanaceae except for a subset of tobacco species have a glycine (**Fig. 3a, Extended Data Fig. 2c and Supplementary Table 1**)[12,22]. Beyond the Solanaceae, this glycine is invariant in angiosperm CLV3 orthologs, is highly conserved in other CLE peptides, and is essential in Arabidopsis CLV3 and CLE41 peptides for precise binding to their receptors (**Extended Data Fig. 2 and Supplementary Table 1**)[12,22,27–30]. These observations suggested that other Solanaceae species with the conserved glycine in their *SlCLE9* orthologs might have more effective ligands, and would also be more potent compensators than tomato *SlCLE9*. We tested this using CRISPR-Cas9 genome editing in groundcherry (**Extended Data Fig. 3**). Notably, null mutation of

groundcherry *pgclv3* resulted in only weak phenotypes similar to petunia *phclv3* mutants (**Fig 3b-e and Extended Data Fig. 3a, b**). We also engineered homozygous *pgcle9* null mutations, which were nearly identical to wild-type (**Fig. 3b-e and Extended Data Fig. 3c**), and consistent with these weak effects, the sizes of primary shoot meristems in both mutants were largely unchanged (**Fig. 3f, g**). Importantly, as in tomato and in petunia, the expression of both *PgCLV3* and *PgCLE9* were upregulated in *pgclv3* meristems (**Fig. 3h, i and Supplementary Table 3**), and *pgclv3 pgcle9* double null mutants were severely fasciated, similar to tomato *slclv3 slcle9* double mutants, confirming conservation of active compensation (**Fig. 3j, k and Extended Data Fig. 3d, e**). Thus, while active compensation is conserved between tomato, petunia, and groundcherry, compensation from *SlCLE9* orthologs in petunia and groundcherry is stronger than in tomato.

Our dissections of active compensation in tomato, petunia, and groundcherry suggested that the conserved glycine of the dodecapeptide is necessary for potent compensation. In further support, two conserved residues (Aspartic acid and Phenylalanine) in SlCLV1, which is the primary receptor of SlCLV3 and SlCLE9 ligands[12], are critical for interaction with the sixth glycine of CLE peptides (**Extended Data Fig. 4**)[29,30]. Solanaceae CLV1 orthologs are invariant in these ligand binding residues (**Extended Data Fig. 4**). To test if the groundcherry and petunia orthologs of CLV1 (PgCLV1 and PhCLV1) are also the primary receptors for PgCLE9 and PhCLE9 as in tomato, we made double mutants between the weakly fasciated groundcherry *pgclv1* and *pgclv3* and also the weakly fasciated petunia *phclv1* and *phclv3* null mutants (**Extended Data Fig. 5)**[31]. Consistently, the double null mutants in both species matched the severe fasciation of groundcherry *pgclv3 pgcle9* double mutants, and importantly, also the tomato *slclv1 slclv3* and *slclv3 slcle9* double mutants (**Fig. 1c, 3j and Extended Data Fig. 5c-e**). These results support the hypothesis that the glycine to serine change in the tomato SlCLE9 dodecapeptide could be reducing binding affinity to SlCLV1, thus explaining weaker compensation in this species.

To test the significance of the glycine, we asked if the genomic sequence of *PgCLE9* (*gPgCLE9*^{PgCLE9}) could complement *slclv3* mutants (**Fig. 4a**). While *slclv3* fasciation is nearly completely suppressed by the genomic sequence of *SlCLV3* (*gSlCLV3*^{SlCLV3}), *gPgCLE9*^{PgCLE9} had no effect (**Fig. 4a, b and Extended Data Fig. 6a, b**). Poor heterologous expression between groundcherry and tomato could explain this result, so we transformed *slclv3* mutants with a construct expressing the groundcherry dodecapeptide from the genomic sequence of tomato *SlCLE9* (*gSlCLE9*^{PgCLE9}) (**Fig. 4a, b and Extended Data Fig. 6a, b**). Surprisingly, this construct

also failed to complement, leading us to ask if strong active compensation depended on the conserved glycine as well as higher expression of dodecapeptides having the glycine. In support of this, in contrast to tomato, the fold-change increases in expression of both groundcherry *PgCLE9* and petunia *PhCLE9* were higher relative to upregulation of *CLV3* in their respective *clv3* mutants (**Fig. 2h, 3i**). As the promoter of tomato *SlCLV3* is more transcriptionally responsive than the promoter of *SlCLE9* to *slclv3* mutations (**Fig. 3h**), we used a construct expressing the groundcherry dodecapeptide from *SlCLV3* genomic sequence (*gSlCLV3*^{PgCLE9}), which strongly suppressed *slclv3* mutants. Notably, this complementation was slightly weaker than with *gSlCLV3*^{SlCLV3}, consistent with active compensation from PgCLE9 and PhCLE9 dodecapeptides in groundcherry and petunia still permitting weak phenotypes of their respective *clv3* mutants (**Fig. 4a, b and Extended Data Fig. 6a, b**). A construct expressing the tomato SlCLE9 dodecapeptide from the same *SlCLV3* genomic sequence (*gSlCLV3*^{SlCLE9}) failed to complement, indicating that higher expression alone is insufficient (**Fig. 4a, b and Extended Data Fig. 6a, b**). Consistently, a weaker expression of PgCLE9 dodecapeptide (*gSlCLE9*^{SlCLE9S6G}) or a stronger expression of SlCLE9 dodecapeptide (*gSlCLV3*^{SlCLE9}*-2*) could only suppress *slclv3 slcle9* double mutants to *slclv3* single mutant phenotypes (**Extended Data Fig. 6c, d)**. Altogether, our results show that changes in both the dodecapeptide and its expression explain evolutionary variation in the strength of compensation between tomato and its relatives groundcherry and petunia (**Fig. 4c**).

Here, we uncovered a dynamic evolution of paralogs interacting in an active compensation relationship. A first step of paralog diversification that can promote their preservation is 'compensatory drift', through which optimal levels of dosage-sensitive genes are maintained by reducing the expression of one paralog and elevating the other[32]. *CLV3* orthologs are dosage-sensitive[33–35], and the consistently higher expression levels of Solanaceae *CLV3* orthologs relative to *SlCLE9* orthologs indicate that compensatory drift and active compensation emerged soon after duplication (**Fig. 2g, 3h**). However, despite this expression rebalancing, we found that *CLV3* compensation degraded multiple times during the Solanaceae family radiation over the last ~30 million years (**Fig. 4d**). At one extreme, *N. benthamiana,* and likely other species that lost their *SlCLE9* orthologs, completely lost active compensation and thus buffering of meristem homeostasis. In tomato, both coding and *cis*-regulatory changes weakened *SlCLE9*, and we pinpointed a critical amino acid change that facilitated partial degradation of compensation from the more potent ancestral state found in groundcherry and petunia (**Fig. 4d**). Thus, the differential

accumulation of genetic variation between *SlCLE9* orthologs in these four Solanaceae species resulted in both qualitative and quantitative differences in compensation potencies. Our finding of extensive coding and *cis*-regulatory variation between *SlCLE9* orthologs suggests a range of potencies could exist in Solanaceae *CLV3* compensation (**Extended Data Fig. 2 and Supplementary Table 1**). For example, even among tobacco species, while *N. benthamiana* lost compensation, *N. obtusifolia* likely has strong compensation due to retention of a glycine-containing *SlCLE9* ortholog, and surprisingly, the sub-genome copies of *SlCLE9* orthologs in *N. attenuata*, *N. tabacum,* and *N. tomentosiformis* each have a glycine and a serine (**Extended Data Fig. 2c and Supplementary Table 1**).

Differences in transcriptional control may play a larger role. Widespread variation in *cis*-regulatory regions among tomato species suggests even greater variation between species in the Solanaceae family[36]. Such diversity, both within and between genera (**Extended Data Fig. 2a**), could result in differences in upregulation of *SlCLE9* orthologs and phenotypes when *CLV3* activity is compromised. Such a wide range of compensation strengths could be a foundation for species-specific phenotypes. Notably, a structural variant that partially disrupts the promoter of *SlCLV3* is a major tomato domestication fruit size QTL, and we found that its severity was mitigated by active compensation from *SlCLE9,* resulting in a more moderate effect that may have facilitated selection[12,37]. The increase in fruit size from this variant may not have emerged if the ancestral version of *SlCLE9* was retained in tomato, and moreover, in groundcherry and other Solanaceae orphan crops with potent *SlCLE9* orthologs, engineering mutations in *CLV3* alone would likely not benefit fruit size[31,38]. Beyond the Solanaceae, variation in meristem shape and form is associated with morphological variation within and between species[39–41]. Such differences could in part be based on variation in compensation between meristem homeostasis genes, which could also influence phenotypic outcomes from engineered variation in CLV network genes[33,35,42].

More broadly, our findings have important implications in understanding and exploiting phenotypic changes caused by natural and engineered variation in other species and gene families. The era of pan-genomes[43–46] continues to uncover remarkable diversity in paralogs, including presence-absence variation, as well as widespread coding and regulatory variation between retained paralogs. Our findings show that such variation, much of which could be cryptic[47–49], can impact phenotypes in unpredictable and subtle ways when members of a gene family are mutated within or between species. Revealing and dissecting diverse paralogous relationships can advance

240 our understanding of how dynamically evolving duplicated genes shape phenotypic variation
241 across short time scales, and improve predictability in trait engineering of both old and new crops.
242

Methods

Plant materials and growth conditions.

Seeds of petunia (*P. hybrida* 'W115', Mitchel diploid) were provided by Prof. Yulong Guo, Southwest University (Chongqing, China). Seeds of tobacco (*N. benthamiana*), groundcherry (*P. grisea*) and tomato (*S. lycopersicum,* cultivar M82) were from Cold Spring Harbor Laboratory (CSHL) seed stocks. All seeds were sown directly in soil and grown in growth chambers, greenhouses or fields at CSHL, New York, USA (tomato, tobacco, groundcherry) and Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, China (petunia). Briefly, groundcherry and tomato seedlings were grown in the greenhouse or field at CSHL as described previously[50]. Tobacco plants were grown under long-day conditions (16 h light, 21°C/8 h dark, 20°C; 40-55% relative humidity; 75 μmol m^{-2} s^{-1}) in the greenhouse at CSHL. *Petunia* plants were grown under long-day conditions (16h light, 25°C/8h dark, 21°C; 50-60% relative humidity; 75 μmol m^{-2} s^{-1}) in growth chambers and greenhouses at Institute of Genetics and Developmental Biology, Chinese Academy of Sciences. All plants were grown under overhead watering (tobacco) or drip irrigation (groundcherry, petunia and tomato), and standard fertilizer regimes.

CRISPR–Cas9 genome editing and plant transformation.

Targeted mutagenesis using the CRISPR-Cas9 system for tobacco, groundcherry, and petunia were performed as described previously[31,51–57]. Briefly, the binary vectors were constructed through Golden Gate cloning as described[51,58], and introduced into tobacco, groundcherry, and petunia by *Agrobacterium tumefaciens*-mediated transformation as described[52,53,57,59]. First-generation transgenic plants were transplanted in soil and genotyped to validate CRISPR-generated mutations by PCR and Sanger sequencing, as previously described[37]. All primer and gRNA sequences are included in **Supplementary Table 4**.

Plant phenotyping and meristem imaging.

All phenotypic quantification data on inflorescences and fruits were performed as previously described[12,37]. Briefly, the phenotypic characterization was performed with biallelic or chimeric T_0 plants (tobacco), and non-transgenic homozygous plants (tobacco, groundcherry, petunia, and tomato) from backcrossing or selfing. CRISPR-generated null mutants of groundcherry and tomato

were sprayed with 400 mgl^{-1} kanamycin, and petunia were sprayed with 100mgl^{-1} kanamycin and genotyped by PCR to verify the absence of the transgenes. We manually counted the floral organs (petal and carpel/locule) from multiple inflorescence and plants. All the exact sample numbers of individual transgenic plants and aggregated organ quantifications are marked in the figures and are collated in the Supplementary Data. Meristem imaging and size quantification were conducted as described previously[37,60]. Briefly, the images of hand-dissected meristems were captured on a Nikon SMZ1500 (tomato), Nikon SMZ25 (groundcherry and tomato). Dissection and stereomicroscope imaging of petunia meristems were carried out under Olympus microscope (SteREO Discovery, v.12).

RNA extraction.

RNA extraction for groundcherry and petunia were conducted as previously described with minor modification[12,50]. Briefly, for total RNA of the groundcherry meristems, the hand-dissected shoot apical meristems were extracted by the ARCTURUS PicoPure RNA Extraction Kit (Applied Biosystems). Three biological replicates were analyzed for groundcherry RNA-seq. 30–35 meristems from groundcherry were collected for each replicate for wild-type and *pgclv3*. Total RNA of the petunia meristems was also extracted by the ARCTURUS PicoPure RNA Extraction Kit (Applied Biosystems). Three biological replicates were examined for petunia RNA-seq. 50–60 meristems from petunia were collected for each replicate for wild-type and *phclv3*.

Meristem transcriptome profiling.

The transcriptome data from tomato meristems were obtained from our previous RNA-seq data deposited in the Sequence Read Archive project (SRP161864) and BioProject (PRJNA491365)[12]. RNA-seq and differentially expressed genes (DEGs) analyses of groundcherry and petunia meristems were performed as previously described with slight modification[12]. Briefly, the libraries for RNA-sequencing (RNA-seq) were prepared by the KAPA mRNA HyperPrep Kit (Roche). The quality of each library was validated with a 2100 Bioanalyzer (Agilent Technologies). Paired-end 75-base sequencing was conducted on the Illumina sequencing platform (NextSeq, Mid-Output). Reads for the wild-type (WT) groundcherry and *pgclv3* mutant were trimmed by quality using Trimmomatic (v.0.32, parameters: ILLUMINACLIP:TruSeq3-PE-2.fa:2:40:15:1:FALSE LEADING:30 TRAILING:30 MINLEN:50)[61] and aligned to the reference transcriptome assembly

of groundcherry[31] for quantification using 'kallisto quant' (v0.46.2, bootstrap: 100)[62]. Kallisto quantification results were used as inputs for 'sleuth' (v0.30.0) in R (v3.5.2) to get normalized estimated counts for each transcript[63]. Expression unit is transcripts per million (TPM) for groundcherry RNA-seq. For RNA-seq of petunia meristems, the libraries were prepared by SMARTer Ultra Low Input RNA for Sequencing Kit (Clontech). The quality of each library was validated with a 2100 Bioanalyzer (Agilent Technologies). Paired-end 150-base sequencing was conducted on the Illumina NovaSeq 6000 sequencing platform (NextSeq, Mid-Output). Reads for the WT petunia and *phclv3* mutant were trimmed by quality using Trimmomatic (v0.36, parameters: ILLUMINACLIP:adapter.fa:2:30:10 LEADING:20 TRAILING:20 SLIDINGWINDOW:4:15 MINLEN:36)[61] and aligned to the reference genome sequence of petunia[64] using hisat2 (v2.1.0) with default parameters[65]. Alignments were sorted with samtools (v1.8)[66] and the RNA-seq reads were assembled using StringTie (v2.0.3) with default parameters[67]. To verify and annotate the transcript of petunia *PhCLE9* (Peaxi162Scf00429:766800-783916), orthologous Blast was performed using tomato *SlCLE9* as a bait and the resulting transcript was confirmed by PCR amplification followed by Sanger sequencing (see **Supplementary Data 9**). The expected read counts and fragments per kilobase of transcript per million mapped reads (FPKM) were also calculated using SringTie (v2.0.3)[67]. The statistical analyses for groundcherry and petunia data were performed in R (v3.5.2) (RStudio (v.1.1.463)) and R (v4.0.3), respectively[68,69]. Significant differential expression between groundcherry WT and *pgclv3* mutant was identified with sleuth (v0.30.0)[63] using q-value ≤ 0.01 cut-offs. Significant differential expression between petunia WT and *phclv3* mutant was confirmed with DESeq2 (v1.30.1)[63,70] using *p-vadue adjusted* (*padj*) ≤ 0.05 and $|\log2\_ratio| \geq 1$.

Transgenic complementation of *PgCLE9, SlCLV3* and *SlCLE9*.

The transgenic lines and genomic DNA sequence for $gSlCLV3^{SlCLV3}$ and $gSlCLV3^{SlCLE9}$ were procured from our previous study[12]. The genomic DNA sequences of *PgCLE9* consisted of $gPgCLE9^{PgCLE9}$ 4471 base pair (bp) in total with 3394 bp upstream, 548 bp of coding sequence containing introns, and 529 bp downstream. The genomic DNA sequences of *SlCLE9* consisted of $gSlCLE9^{SlCLE9}$ 4140 bp in total with 3263 bp upstream, 403 bp of coding sequence containing introns, and 474 bp downstream. Site-directed mutageneses were performed to substitute the SlCLE9 dodecapeptide into PgCLE9 within $gSlCLE9^{SlCLE9}$ ($gSlCLE9^{PgCLE9}$) and the SlCLV3

dodecapeptide into PgCLE9 within *gSlCLV3*^{SlCLV3} (*gSlCLV3*^{PgCLE9}). The PCR products were amplified from the vectors including the genomic region of *SlCLV3* (pICH47742-*gSlCLV3*^{SlCLV3}) and *SlCLE9* (pICH47742- *gSlCLE9*^{SlCLE9}) with overlapping primers (**Supplementary Table 4**) using KOD One^{TM} PCR Master Mix (TOYOBO). Then, the amplified PCR products were digested using DpnI (New England Biolabs) and transformed into DH5a competent cells. The sequences of the resulting plasmids were confirmed by Sanger sequencing with multiple primers (**Supplementary Table 4**). The Level 1 vectors (pICH47742-*gPgCLE9*^{PgCLE9}, *gSlCLE9*^{PgCLE9} and *gSlCLV3*^{PgCLE9}) were assembled with the construct pICH47732-NOSpro::NPTII into the binary vector pICSL4723 through Golden Gate cloning as previously described[51,58,71]. The binary vectors were introduced into the tomato *slclv3* mutant by *Agrobacterium tumefaciens*-mediated transformation as previously described[53]. The genomic DNA sequences of *SlCLV3* consisted of *gSlCLV3*^{SlCLV3}*-2* 3213 bp in total with 1995 bp upstream, 600 bp of coding sequence containing introns, and 618 bp downstream. The genomic DNA sequences of *SlCLE9* consisted of *gSlCLE9*^{SlCLE9}*-2* 2740 bp in total with 1996 bp upstream, 403 bp of coding sequence containing introns, and 341 bp downstream. Site-directed mutagenesis was performed to substitute the SlCLV3 dodecapeptide into SlCLE9 within *gSlCLV3*^{SlCLV3} (*gSlCLV3*^{SlCLE9}*-2*) and the SlCLE9 dodecapeptide into SlCLE9^{S6G} within *gSlCLE9*^{SlCLE9} (*gSlCLE9*^{SlCLE9S6G}). The PCR products were amplified from the vectors including the genomic region of *SlCLV3* (pDONOR221-*gSlCLV3*^{SlCLV3}*-2*) and *SlCLE9* (pDONOR221- *gSlCLE9*^{SlCLE9}*-2*) with overlapping primers (**Supplementary Table 4**) using KOD One^{TM} PCR Master Mix (TOYOBO). Then, the amplified PCR products were digested using DpnI (New England Biolabs) and transformed into DH5a competent cells. The sequences of the resulting plasmids were confirmed by Sanger sequencing with multiple primers (**Supplementary Table 4**), and colonies were recombined into binary vector pGWB401[72] for transgenic complementation. The binary vectors were introduced into the tomato *slclv3 slcle9* double mutant by *Agrobacterium tumefaciens*-mediated transformation as previously described[53]. Transgenic lines were confirmed by PCR and kanamycin resistance, and at least three independent transgenic lines from each construct were used for data collection (see **Supplementary Data**).

Conserved noncoding sequence (CNS) analysis.

Analysis of conserved non-coding sequences (CNSs) is a common approach to identify putative *cis*-regulatory sequences of genes (e.g. promoters, enhancers). Solanaceae orthologous genes of

SlCLV3 and *SlCLE9* for synteny analysis and CNSs in the promoter regions surrounding the orthologs of *SlCLV3* and *SlCLE9* were identified using our previously developed Conservatory algorithm (v1.0), using default parameters[73]. In parallel, all of the genomes were scanned with tBLASTn to find mis- or unannotated protein coding regions for each gene. CNSs in the promoter regions were called by Conservatory using default parameters[73]. To calculate protein identity percentages and dodecapeptide identity percentages, protein sequences were aligned by MAFFT (v.7.45) using BLOSUM62 matrix and 'E-INS-i' and 'G-INS-i' algorithm respectively[74].

Statistical analyses.

Statistical calculations were conducted using R(v3.5.2 and v4.0.3)[68] and Microsoft Excel, as previously described[50]. Statistical analyses were performed using a two-tailed, two-sample *t*-test and a one-way analysis of variance (ANOVA) with Tukey test. The exact sample sizes (n) and all raw data for each experimental group/condition are given as discrete numbers in each figure panel and Supplementary Data. Additional information is available in the Nature Research Reporting Summary, which includes statements on statistics, software used and data availability.

Data availability

Raw data and information for CRISPR-generated alleles, all quantifications, synteny analysis, and exact *P* values (One-way ANOVA and Tukey test) are in **Supplementary Data**. The raw Sanger sequence traces for edited sequences are in **Supplementary Data 8**. The groundcherry and petunia BioProject accession numbers are PRJNA704671 and PRJNA750419, respectively.

Acknowledgements

We thank members of the Lippman laboratory for comments and discussions, and also critical friends Y. Eshed and M. Bartlett. We thank S. Soyk for discussions and initial peptide sequence analysis in tomato, and J. Kim and A. Krainer for technical support. We thank A. Horowitz Doyle, K. Swartwood, M. Tjahjadi, L. Randall, and P. Keen from the Van Eck lab for performing tobacco, groundcherry and tomato transformations. We thank T. Mulligan, K. Schlecht, A. Krainer and S. Qiao for assistance with plant care. This research was supported by National Natural Science Foundation of China (grant no. 31972423 and 31991183) and Chinese Academy of Sciences (grant no. 153E11KYSB20180019) to C.X., the Howard Hughes Medical Institute, an Agriculture and

Food Research Initiative competitive grant from the USDA National Institute of Food and Agriculture (grant no. 2016-67013-24452) and the National Science Foundation Plant Genome Research Program (grant no. IOS-1732253 and IOS-1546837) to Z.B.L.

Author contributions

C.-T.K. designed the research and conducted the experiments, prepared the figures and wrote the manuscript. L.T. performed the petunia CRISPR experiments, tomato transgenic complementation tests, genetic, RNA-seq and phenotypic analyses. X.W. performed the groundcherry RNA-seq, phenotypic analyses, and wrote the manuscript. I.G performed the tobacco genetic and phenotypic analyses. A.H. performed CLE family analyses. G.R. characterized CRISPR mutations. J.V.E generated transgenic plants and CRISPR lines. C.X. supervised and led the petunia CRISPR experiments and tomato transgenic complementation tests, genetic, RNA-seq and phenotypic analyses, contributed ideas and edited the manuscript. Z.B.L. conceived and led the research, supervised and performed the experiments, prepared the figures and wrote the manuscript. All authors read, edited, and approved the manuscript.

Competing interests

The authors declare that they have no competing interests.

Figure legends

Fig. 1. Loss of the tobacco *SlCLE9* ortholog abolished compensation.

a, Shoot and inflorescence of tomato wild-type (WT). White arrowheads, inflorescences. **b**, Shoot and inflorescence of tomato *slclv3*. White arrowheads, inflorescences; red arrowheads, branches. **c**, Side and top-down view of tomato *slclv3 slcle9* shoot, inflorescence/floral meristem, and primary inflorescence. The red arrowhead indicates a fasciated shoot stem. **d**, Presence-absence variation of *SlCLE9* orthologs in the Solanaceae. The blue checkmarks and the red Xs indicate presence and absence of the orthologs, respectively. **e**, Gene structures, and CRISPR-generated mutations of *NbCLV3a and NbCLV3b*. Orange rectangles indicate the CLE dodecapeptides regions. Targeted gRNA and protospacer-adjacent motif (PAM) sequences are highlighted in red and bold underlined, respectively. Blue letters and dashes indicate insertions and deletions, respectively. Numbers in parentheses represent gap lengths. DNA sequences of gRNA target site 2 for both *NbCLV3a* and *NbCLV3b* are identical. **f**, Shoot, flower, and fruit pod of tobacco WT. White arrowheads, flowers. **g**, Side and top-down views of *nbclv3a/b* null mutants showing the shoot and primary flower. Red arrowheads indicate fasciated primary shoot (left panel) and shoot branches (right panel). **h**, Sepal number of primary flower from tobacco WT, WT sibling plants (WT sibs) and *nbclv3a/b* plants. **i**, Branch number of WT, WT sibs and *nbclv3a/b*. **j**, Primary shoot apical meristems from WT and *nbclv3a/b*. Red dotted lines mark width and height for meristem size quantification. 7L, 7th leaf primordium. **k**, Quantification of meristem width and height from WT, WT sibs and *nbclv3a/b*. Box plots, 25th-75th percentile; center line, median; whiskers, full data range in **h**, **i** and **k**. Exact sample sizes (n) for replicate types are indicated in **h**, **i** and **k**. Letters indicate significance groups at $P < 0.01$ (One-way ANOVA and Tukey test) in **h**, **i** and **k**. Different letters between genotypes indicate significance in **h**, **i** and **k** (See Supplementary Data 7 for specific *P* values). WT sibs are a mix of *nbclv3b* and *nbclv3a/+ nbclv3b* genotypes, which show wild-type phenotypes in **h**, **i** and **k** (See Supplementary Data 3). At least twice experiments were repeated independently with similar results.

Fig. 2. Weak fasciation of *phclv3* mutants in petunia indicates more potent compensation.

a, Gene structure and sequences of two *phclv3* null alleles. Guide RNA and PAM sequences are highlighted in red and bold underlined, respectively. The orange rectangles in the gene structures

represent the regions for CLE dodecapeptides. Numbers in parentheses represent gap lengths. Blue dashes indicate deletions. **b**, Shoot of petunia WT and *phclv3* plants. White arrowheads, flowers. **c**, Representative flowers and fruit pods of petunia WT and *phclv3* plants. Red arrowheads mark petals or carpels. Percentages indicate the proportions of flower and pod phenotypes. **d**, Quantification of petal and carpel numbers of WT and *phclv3*. **e**, Primary shoot apical meristems from petunia WT and *phclv3*. Red dotted lines mark width and height for meristem size quantification. 22L, 22th leaf primordium. **f**, Quantification of meristem width and height from petunia WT and *phclv3*. **g**, Normalized read counts of *PhCLV3* and *PhCLE9* from WT and *phclv3* meristems. **h**, Expression fold-change of *PhCLV3* and *PhCLE9* relative to the normalized counts of WT from *phclv3*. Box plots, 25th-75th percentile; center line, median; whiskers, full data range in **d**, **f**, **g** and **h**. *P* values (two-tailed, two-sample *t*-test) in **d**, **f**, **g** and **h**. Exact sample sizes (n) are shown as discrete numbers in **d**, **f**, **g** and **h**. Each replicate (n) is from 50-60 meristems in **g** and **h**. At least twice experiments were repeated independently with similar results.

Fig. 3. A highly conserved dodecapeptide amino acid is associated with potent compensation in groundcherry.

a, CLE protein structure and dodecapeptide sequences of *SlCLE9* and *SlCLV3* orthologs in the Solanaceae. **b**, Shoot and inflorescences of groundcherry WT, *pgclv3* and *pgcle9* plants. Red arrowheads mark two side shoots that develop after single-flowered inflorescences. **c**, Representative flowers and fruits from groundcherry WT, *pgclv3,* and *pgcle9* plants. Scale bar, 1 cm. **d**, Representative flowers and fruits from tomato WT, *slclv3,* and *slcle9* plants. White arrowheads mark petals or locules. Scale bar, 1 cm. **e**, Quantification of petal and locule numbers from groundcherry WT, *pgclv3, pgcle9* and tomato WT, *slclv3,* and *slcle9* plants. **f**, Primary shoot apical meristems from groundcherry WT, *pgclv3, pgcle9* and tomato WT, *slclv3,* and *slcle9* plants. 7L, 8L: 7th and 8th leaf primordia, respectively. Red dotted lines indicate width and height for meristem size measurements, Scale bar, 200 μm. **g**, Quantification of meristem width and height from groundcherry WT, *pgclv3, pgcle9*, tomato WT, *slclv3,* and *slcle9* plants. **h**, Normalized RNA-seq read counts of *SlCLV3*, *SlCLE9*, *PgCLV3*, and *PgCLE9* from tomato WT, *slclv3,* groundcherry WT and *pgclv3* meristems. **i**, Expression fold-change of *SlCLV3*, *SlCLE9*, *PgCLV3*, and *PgCLE9* relative to the normalized counts of WT expression of these genes in the indicated genotypes. **j**, Side and top-down views of a *pgclv3 pgcle9* double mutant shoot, inflorescence/floral meristem,

and primary flower. Red arrowheads indicate branches that emerged after the primary flower. **k**, Branch number of WT, *pgclv3*, *pgcle9*, and *pgclv3 pgcle9* plants. Box plots, 25^{th}-75^{th} percentile; center line, median; whiskers, full data range in **e**, **g**, **h**, **i** and **k**. The letters indicate the significance groups at $P < 0.01$ (One-way ANOVA and Tukey test) in **e**, **g** and **k**. Different letters between genotypes indicate significance in **e**, **g** and **k** (See Supplementary Data 7 for specific *P* values). *P* values (two-tailed, two-sample *t*-test) in **h** and **i**. Exact sample sizes (n) are shown in **e**, **g**, **h**, **i** and **k**. Each replicate (n) is from 30-35 meristems in **h** and **i**. At least twice experiments were repeated independently with similar results.

Fig. 4. Variation in Solanaceae compensation is due to changes in both the SlCLE9 ortholog dodecapeptide and its expression.

a, Diagrams of constructs used for complementation tests. $gPgCLE9^{PgCLE9}$ (*PgCLE9* genomic DNA). $gSlCLE9^{PgCLE9}$ (*SlCLE9* genomic DNA including the sequence for PgCLE9 dodecapeptide). $gSlCLV3^{PgCLE9}$ (*SlCLV3* genomic DNA including the sequence for PgCLE9 dodecapeptide). $gSlCLV3^{SlCLE9}$ (*SlCLV3* genomic DNA including the sequence for SlCLE9 dodecapeptide). Black and orange rectangles mark the coding sequences and the dodecapeptide sequences, respectively. The numbers with minus (-) and plus (+) signs indicate the positions of the upstream sequences and the downstream sequences from the adenines of start codons, respectively. **b**, Locule number quantification from WT and *slclv3* mutants compared to T_1 transgenic plants of $gSlCLV3^{SlCLV3}$, $gPgCLE9^{PgCLE9}$, $gSlCLE9^{PgCLE9}$, $gSlCLV3^{PgCLE9}$, and $gSlCLV3^{SlCLE9}$. Box plots, 25^{th}-75^{th} percentile; center line, median; whiskers, full data range. The letters indicate the significance groups at $P < 0.01$ (One-way ANOVA and Tukey test). Different letters between genotypes indicate significance (See Supplementary Data 7 for specific *P* values). Exact sample sizes (n) are shown as discrete numbers. Data are based on at least 10 independent transgenic lines for each construct. At least twice experiments were repeated independently with similar results. **c**, A proposed model for differences in active compensation between tomato and groundcherry. The more potent active compensation in groundcherry compared to tomato is due to both the glycine-containing PgCLE9 dodecapeptide and its higher expression. **d,** Summary and model of the dynamic evolution of *SlCLV3* and *SlCLE9* orthologs and their compensation relationships in the Solanaceae. Dark blue, blue, and sky blue rectangles indicate the coding region of the genes. Arrows and their thickness represent gene expression and their relative levels, respectively. Numbers above the arrows

indicate hypothetical relative proportions of *SlCLV3* and *SlCLE9* ortholog expression levels. 'G' and 'S' within the rectangles denote the sixth amino acid of each CLE dodecapeptide. Dashed rectangles mark deletions of the coding region, resulting in pseudogenes (pepper and tobacco) and complete gene loss (eggplant, potato) in each genome. The red gradient bar reflects the loss of active compensation and its degree, depending on the indicated genetic variation.

References

1. Conrad, B. & Antonarakis, S. E. Gene Duplication: A Drive for Phenotypic Diversity and Cause of Human Disease. *Annu. Rev. Genomics Hum. Genet.* **8**, 17–35 (2007).
2. Carretero-Paulet, L. & Fares, M. A. Evolutionary Dynamics and Functional Specialization of Plant Paralogs Formed by Whole and Small-Scale Genome Duplications. *Mol. Biol. Evol.* **29**, 3541–3551 (2012).
3. Panchy, N., Lehti-Shiu, M. & Shiu, S.-H. Evolution of Gene Duplication in Plants1[OPEN]. *Plant Physiol.* **171**, 2294–2316 (2016).
4. Copley, S. D. Evolution of new enzymes by gene duplication and divergence. *FEBS J.* **287**, 1262–1283 (2020).
5. Kuzmin, E., Taylor, J. S. & Boone, C. Retention of duplicated genes in evolution. *Trends Genet.* **38**, 59–72 (2022).
6. Diss, G., Ascencio, D., DeLuna, A. & Landry, C. R. Molecular mechanisms of paralogous compensation and the robustness of cellular networks. *J. Exp. Zoolog. B Mol. Dev. Evol.* **322**, 488–499 (2014).
7. Kafri, R., Springer, M. & Pilpel, Y. Genetic Redundancy: New Tricks for Old Genes. *Cell* **136**, 389–392 (2009).
8. DeLuna, A., Springer, M., Kirschner, M. W. & Kishony, R. Need-Based Up-Regulation of Protein Levels in Response to Deletion of Their Duplicate Genes. *PLOS Biol.* **8**, e1000347 (2010).
9. Somssich, M., Je, B. I., Simon, R. & Jackson, D. CLAVATA-WUSCHEL signaling in the shoot meristem. *Development* **143**, 3238–3248 (2016).
10. Fletcher, J. C. The CLV-WUS Stem Cell Signaling Pathway: A Roadmap to Crop Yield Optimization. *Plants* **7**, 87 (2018).

11. Soyars, C. L., James, S. R. & Nimchuk, Z. L. Ready, aim, shoot: stem cell regulation of the shoot apical meristem. *Curr. Opin. Plant Biol.* **29**, 163–168 (2016).

12. Rodriguez-Leal, D. *et al.* Evolution of buffering in a genetic circuit controlling plant stem cell proliferation. *Nat. Genet.* **51**, 786–792 (2019).

13. Jiao, W.-B. *et al.* The Evolutionary Dynamics of Genetic Incompatibilities Introduced by Duplicated Genes in Arabidopsis thaliana. *Mol. Biol. Evol.* **38**, 1225–1240 (2021).

14. Innan, H. & Kondrashov, F. The evolution of gene duplications: classifying and distinguishing between models. *Nat. Rev. Genet.* **11**, 97–108 (2010).

15. Albalat, R. & Cañestro, C. Evolution by gene loss. *Nat. Rev. Genet.* **17**, 379–391 (2016).

16. El-Brolosy, M. A. & Stainier, D. Y. R. Genetic compensation: A phenomenon in search of mechanisms. *PLOS Genet.* **13**, e1006780 (2017).

17. Hanada, K. *et al.* Functional Compensation of Primary and Secondary Metabolites by Duplicate Genes in Arabidopsis thaliana. *Mol. Biol. Evol.* **28**, 377–382 (2011).

18. Kafri, R., Bar-Even, A. & Pilpel, Y. Transcription control reprogramming in genetic backup circuits. *Nat. Genet.* **37**, 295–299 (2005).

19. DeLuna, A. *et al.* Exposing the fitness contribution of duplicated genes. *Nat. Genet.* **40**, 676–681 (2008).

20. Freeling, M. Bias in Plant Gene Content Following Different Sorts of Duplication: Tandem, Whole-Genome, Segmental, or by Transposition. *Annu. Rev. Plant Biol.* **60**, 433–453 (2009).

21. Vavouri, T., Semple, J. I. & Lehner, B. Widespread conservation of genetic redundancy during a billion years of eukaryotic evolution. *Trends Genet.* **24**, 485–488 (2008).

562 22. Goad, D. M., Zhu, C. & Kellogg, E. A. Comprehensive identification and clustering of
563 CLV3/ESR-related (CLE) genes in plants finds groups with potentially shared function. *New*
564 *Phytol.* **216**, 605–616 (2017).
565 23. Fletcher, J. C. Recent Advances in Arabidopsis CLE Peptide Signaling. *Trends Plant Sci.*
566 **25**, 1005–1016 (2020).
567 24. Hirakawa, Y. & Sawa, S. Diverse function of plant peptide hormones in local signaling
568 and development. *Curr. Opin. Plant Biol.* **51**, 81–87 (2019).
569 25. Whitewoods, C. D. *et al.* CLAVATA Was a Genetic Novelty for the Morphological
570 Innovation of 3D Growth in Land Plants. *Curr. Biol.* **28**, 2365-2376.e5 (2018).
571 26. Stuttmann, J. *et al.* Highly efficient multiplex editing: one-shot generation of 8×
572 Nicotiana benthamiana and 12× Arabidopsis mutants. *Plant J.* **106**, 8–22 (2021).
573 27. Fiers, M. *et al.* The CLAVATA3/ESR Motif of CLAVATA3 Is Functionally Independent
574 from the Nonconserved Flanking Sequences. *Plant Physiol.* **141**, 1284–1292 (2006).
575 28. Ogawa, M., Shinohara, H., Sakagami, Y. & Matsubayashi, Y. Arabidopsis CLV3 Peptide
576 Directly Binds CLV1 Ectodomain. *Science* **319**, 294–294 (2008).
577 29. Zhang, H., Lin, X., Han, Z., Qu, L.-J. & Chai, J. Crystal structure of PXY-TDIF complex
578 reveals a conserved recognition mechanism among CLE peptide-receptor pairs. *Cell Res.* **26**,
579 543–555 (2016).
580 30. Li, Z., Chakraborty, S. & Xu, G. Differential CLE peptide perception by plant receptors
581 implicated from structural and functional analyses of TDIF-TDR interactions. *PLOS ONE* **12**,
582 e0175317 (2017).
583 31. Lemmon, Z. H. *et al.* Rapid improvement of domestication traits in an orphan crop by
584 genome editing. *Nat. Plants* **4**, 766–770 (2018).

585 32. Thompson, A., Zakon, H. H. & Kirkpatrick, M. Compensatory Drift and the Evolutionary
586 Dynamics of Dosage-Sensitive Duplicate Genes. *Genetics* **202**, 765–774 (2016).
587 33. Rodríguez-Leal, D., Lemmon, Z. H., Man, J., Bartlett, M. E. & Lippman, Z. B.
588 Engineering Quantitative Trait Variation for Crop Improvement by Genome Editing. *Cell* **171**,
589 470-480.e8 (2017).
590 34. Wang, X. *et al.* Dissecting cis-regulatory control of quantitative trait variation in a plant
591 stem cell circuit. *Nat. Plants* **7**, 419–427 (2021).
592 35. Liu, L. *et al.* Enhancing grain-yield-related traits by CRISPR–Cas9 promoter editing of
593 maize CLE genes. *Nat. Plants* **7**, 287–294 (2021).
594 36. Alonge, M. *et al.* Major Impacts of Widespread Structural Variation on Gene Expression
595 and Crop Improvement in Tomato. *Cell* **182**, 145-161.e23 (2020).
596 37. Xu, C. *et al.* A cascade of arabinosyltransferases controls shoot meristem size in tomato.
597 *Nat. Genet.* **47**, 784–792 (2015).
598 38. Samuels, J. Biodiversity of Food Species of the Solanaceae Family: A Preliminary
599 Taxonomic Inventory of Subfamily Solanoideae. *Resources* **4**, 277–322 (2015).
600 39. Leiboff, S. *et al.* Genetic control of morphometric diversity in the maize shoot apical
601 meristem. *Nat. Commun.* **6**, 8974 (2015).
602 40. Leiboff, S., DeAllie, C. K. & Scanlon, M. J. Modeling the Morphometric Evolution of
603 the Maize Shoot Apical Meristem. *Front. Plant Sci.* **7**, 1651 (2016).
604 41. Schnablová, R., Herben, T. & Klimešová, J. Shoot apical meristem and plant body
605 organization: a cross-species comparative study. *Ann. Bot.* **120**, 833–843 (2017).
606 42. Yang, Y. *et al.* Precise editing of CLAVATA genes in Brassica napus L. regulates
607 multilocular silique development. *Plant Biotechnol. J.* **16**, 1322–1335 (2018).

608 43. Khan, A. W. *et al.* Super-Pangenome by Integrating the Wild Side of a Species for
609 Accelerated Crop Improvement. *Trends Plant Sci.* **25**, 148–158 (2020).
610 44. Della Coletta, R., Qiu, Y., Ou, S., Hufford, M. B. & Hirsch, C. N. How the pan-genome
611 is changing crop genomics and improvement. *Genome Biol.* **22**, 3 (2021).
612 45. Bayer, P. E., Golicz, A. A., Scheben, A., Batley, J. & Edwards, D. Plant pan-genomes are
613 the new reference. *Nat. Plants* **6**, 914–920 (2020).
614 46. Sherman, R. M. & Salzberg, S. L. Pan-genomics in the human genome era. *Nat. Rev.*
615 *Genet.* **21**, 243–254 (2020).
616 47. Paaby, A. B. & Rockman, M. V. Cryptic genetic variation: evolution's hidden substrate.
617 *Nat. Rev. Genet.* **15**, 247–258 (2014).
618 48. Wagner, A. The molecular origins of evolutionary innovations. *Trends Genet.* **27**, 397–
619 410 (2011).
620 49. Zheng, J., Payne, J. L. & Wagner, A. Cryptic genetic variation accelerates evolution by
621 opening access to diverse adaptive peaks. *Science* **365**, 347–353 (2019).
622 50. Kwon, C.-T. *et al.* Rapid customization of Solanaceae fruit crops for urban agriculture.
623 *Nat. Biotechnol.* **38**, 182–188 (2020).
624 51. Brooks, C., Nekrasov, V., Lippman, Z. B. & Van Eck, J. Efficient gene editing in tomato
625 in the first generation using the clustered regularly interspaced short palindromic
626 repeats/CRISPR-associated9 system. *Plant Physiol.* **166**, 1292–1297 (2014).
627 52. Swartwood, K. & Van Eck, J. Development of plant regeneration and Agrobacterium
628 tumefaciens-mediated transformation methodology for Physalis pruinosa. *Plant Cell Tissue*
629 *Organ Cult. PCTOC* **137**, 465–472 (2019).

630 53. Van Eck, J., Keen, P. & Tjahjadi, M. Agrobacterium tumefaciens-Mediated
631 Transformation of Tomato. in *Transgenic Plants: Methods and Protocols* (eds. Kumar, S.,
632 Barone, P. & Smith, M.) 225–234 (Springer New York, 2019). doi:10.1007/978-1-4939-8778-
633 8_16.
634 54. Zhang, B., Yang, X., Yang, C., Li, M. & Guo, Y. Exploiting the CRISPR/Cas9 System
635 for Targeted Genome Mutagenesis in Petunia. *Sci. Rep.* **6**, 20315 (2016).
636 55. Zhang, B. *et al.* CRISPR/Cas9-mediated targeted mutation reveals a role for AN4 rather
637 than DPL in regulating venation formation in the corolla tube of Petunia hybrida. *Hortic. Res.*
638 **8**, 1–9 (2021).
639 56. Jiang, W. *et al.* Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene
640 modification in Arabidopsis, tobacco, sorghum and rice. *Nucleic Acids Res.* **41**, e188–e188
641 (2013).
642 57. Gantner, J., Ordon, J., Kretschmer, C., Guerois, R. & Stuttmann, J. An EDS1-SAG101
643 Complex Is Essential for TNL-Mediated Immunity in Nicotiana benthamiana. *Plant Cell* **31**,
644 2456–2474 (2019).
645 58. Werner, S., Engler, C., Weber, E., Gruetzner, R. & Marillonnet, S. Fast track assembly of
646 multigene constructs using Golden Gate cloning and the MoClo system. *Bioeng. Bugs* **3**, 38–
647 43 (2012).
648 59. van der Meer, I. M. Agrobacterium-Mediated Transformation of Petunia Leaf Discs. in
649 *Plant Cell Culture Protocols* (eds. Loyola-Vargas, V. M. & Vázquez-Flota, F.) 265–272
650 (Humana Press, 2006). doi:10.1385/1-59259-959-1:265.

60. Park, S. J., Jiang, K., Schatz, M. C. & Lippman, Z. B. Rate of meristem maturation determines inflorescence architecture in tomato. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 639–644 (2012).

61. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).

62. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527 (2016).

63. Pimentel, H., Bray, N. L., Puente, S., Melsted, P. & Pachter, L. Differential analysis of RNA-seq incorporating quantification uncertainty. *Nat. Methods* **14**, 687–690 (2017).

64. Bombarely, A. *et al.* Insight into the evolution of the Solanaceae from the parental genomes of Petunia hybrida. *Nat. Plants* **2**, 1–9 (2016).

65. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357–360 (2015).

66. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

67. Pertea, M. *et al.* StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **33**, 290–295 (2015).

68. R Core Team (2015). *R: A language and environment for statistical computing.* (R Found. Stat. Comput. Vienna, Austria.).

69. LoVerso, P. R. & Cui, F. A Computational Pipeline for Cross-Species Analysis of RNA-seq Data Using R and Bioconductor. *Bioinforma. Biol. Insights* **9**, 165–174 (2015).

70. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

71. Weber, E., Engler, C., Gruetzner, R., Werner, S. & Marillonnet, S. A Modular Cloning System for Standardized Assembly of Multigene Constructs. *PLoS ONE* **6**, e16765 (2011).

72. Nakagawa, T. *et al.* Improved Gateway binary vectors: high-performance vectors for creation of fusion constructs in transgenic analysis of plants. *Biosci. Biotechnol. Biochem.* **71**, 2095–2100 (2007).

73. Hendelman, A. *et al.* Conserved pleiotropy of an ancient plant homeobox gene uncovered by cis-regulatory dissection. *Cell* **184**, 1724-1739.e16 (2021).

74. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).

Extended Data Figure legends

Extended Data Fig. 1. CRISPR-generated mutations of the tobacco *NbCLV3a* and *NbCLV3b* genes.

a, CRISPR-generated sequences of *nbclv3a* mutant alleles. **b**, CRISPR-generated sequences of *nbclv3b* mutant alleles. Guide RNA and PAM sequences are highlighted in red and bold underlined, respectively. Blue letters and dashes indicate insertions and deletions, respectively. Numbers in parentheses represent gap lengths. **c**, Shoots and inflorescences of *nbclv3a/b* T_0 plants. Three strong lines (*nbclv3a/b*$^{CR\text{-}3\text{-}T0}$, *nbclv3a/b*$^{CR\text{-}4\text{-}T0}$ and *nbclv3a/b*$^{CR\text{-}5\text{-}T0}$) show similar phenotypes compared to null *nbclv3a/b* mutants in **Fig. 1g**. Weak (*nbclv3a/b*$^{CR\text{-}6\text{-}T0}$) and moderate (*nbclv3a/b*$^{CR\text{-}7\text{-}T0}$) lines show regular shoot architecture but fasciated floral organs. White arrowheads indicate flowers. **d**, Sepal number of weak and moderate *nbclv3a/b* T_0 plants. Box plots, 25th-75th percentile; center line, median; whiskers, full data range. The letters indicate the significance groups at $P < 0.01$ (One-way ANOVA and Tukey test). Different letters between genotypes indicate significance (See Supplementary Data 7 for specific P values). The exact sample sizes (n) are shown as discrete numbers. At least twice experiments were repeated independently with similar results.

Extended Data Fig. 2. Conserved noncoding sequence (CNS) analysis of the promoter regions of *SlCLV3* and *SlCLE9* orthologs in the Solanaceae family.

a, Conservatory analysis of Solanaceae *CLV3* and *CLE9* promoters. Purple boxes define highly similar regions of each gene's orthologs in the indicated species, and dark purple boxes define highly similar regions of the paralogous gene (e.g. *CLV3B* or *CLE9B*) in the indicated species. Green boxes define Solanaceae CNSs. **b**, Multiple alignment and logo sequences of SlCLV3 dodecapeptide orthologs in the Solanaceae family. **c**, Multiple alignment and logo sequences of SlCLE9 dodecapeptide orthologs in the Solanaceae family.

Extended Data Fig. 3. CRISPR-generated mutations of groundcherry *PgCLV3* and *PgCLE9*.

a, Gene structure and sequences of *pgclv3* CRISPR mutants. **b**, Flowers and fruits of *pgclv3*. White arrowheads mark petals or locules. Percentages indicate the proportions of flower and fruit phenotypes. Scale bar, 1 cm. **c**, Gene structure and sequences of *pgcle9* CRISPR mutants. The

orange rectangles in the gene structures indicate the regions of the CLE dodecapeptides in **a** and **c**. Guide RNA and PAM sequences are highlighted in red and bold underlined, respectively, in **a** and **c**. Blue letters and dashes indicate insertions and deletions, respectively, in **a** and **c**. Numbers in parentheses represent gap lengths in **a** and **c**. **d**, Shoot and an extremely fasciated primary flower of the *pgclv3 pgcle9* double mutant. **e**, Development of extra shoots (S) from the primary shoot and apex of a *pgclv3 pgcle9* double mutant. L, leaf petioles. At least twice experiments were repeated independently with similar results.

Extended Data Fig. 4. Sequence alignments of CLV1 receptor homologs.

a, Alignment of the Solanaceae CLV1 protein sequences. Red letters indicate the two ultra-conserved amino acids involved in the physical binding of CLE dodecapeptides. **b**, Alignment of CLV1 homologs in angiosperms. All the sequences are from the Phytozome v12.1 database (phytozome.jgi.doe.gov). Yellow highlights mark the conserved Asp and Phe. Detailed sequence information is shown in Supplementary Data 10.

Extended Data Fig. 5. Groundcherry *pgclv1 pgclv3* and petunia *phclv1 phclv3* double mutants are severely fasciated like tomato *slclv1 slclv3* double mutants.

a, Gene structure and sequences of two *phclv1* CRISPR mutants. Guide RNA and PAM sequences are highlighted in red and bold underlined, respectively. Blue letters and dashes indicate insertions and deletions, respectively. Numbers in parentheses represent gap lengths. **b**, Flowers, fruits/pods, and meristems from *pgclv1*, *phclv1*, and *slclv1* single mutants. White arrowheads mark petals or locules. 7L, 7th leaf primordium, 8L, 8th leaf primordium. 22L, 22th leaf primordium. **C**, Side and top-down views of a *pgclv1 pgclv3* double mutant shoot, inflorescence/floral meristem, and primary inflorescence. 6L, 6th leaf primordium. **D**, Side and top-down views of a *phclv1 phclv3* double mutant shoot and primary flower. **E**, Side and top-down views of a *slclv1 slclv3* double mutant shoot, flower, vegetative meristem and primary inflorescence. Fasciated flowers and vegetative meristems are shown in insets of **c** and **e**. **f**, **g**, Petal (**f**) and locule (**g**) numbers of groundcherry WT, *pgclv1*, *pgclv1 pgclv3, pgclv1 pgcle9*, and petunia WT, *phclv1*, and tomato WT, *slclv1*, *slclv1 slclv3*, and *slclv1 slcle9*. Not that all three Solanaceae *clv1* single mutant fasciation phenotypes are similarly weak. Box plots, 25th-75th percentile; center line, median; whiskers, full data range in **d** and **e**. The letters indicate the significance groups at $P < 0.01$ (One-way ANOVA

and Tukey test) in **f** and **g**. Different letters between genotypes indicate significance in **f** and **g** (See Supplementary Data 7 for specific *P* values). *P* values (two-tailed, two-sample *t*-test) in **f** and **g**. Exact sample sizes (n) are shown in **f** and **g**. At least twice experiments were repeated independently with similar results.

Extended Data Fig. 6. Transgenic complementation tests of tomato *slclv3* single and *slclv3 slcle9* double mutants.

a, **b**, Complementation tests of tomato *slclv3* single mutants. Inflorescence images (**a**) and petal number quantifications (**b**) of WT and *slclv3* compared to the T_1 transgenic plants $gSlCLV3^{SlCLV3}$, $gPgCLE9^{PgCLE9}$, $gSlCLE9^{PgCLE9}$, $gSlCLV3^{PgCLE9}$, and $gSlCLV3^{SlCLE9}$. **c**, Diagrams of the constructs used for complementation tests of *slclv3 slcle9* double mutants. $gSlCLV3^{SlCLV3}$ (*SlCLV3* genomic DNA). $gSlCLV3^{SlCLE9}$ (*SlCLV3* genomic DNA including the sequence for SlCLE9 dodecapeptide). $gSlCLE9^{SlCLE9}$ (*SlCLE9* genomic DNA). $gSlCLE9^{SlCLE9S6G}$ (*SlCLE9* genomic DNA including the sequence for PgCLE9 dodecapeptide). Black and orange rectangles mark the coding sequences and the dodecapeptide sequences, respectively. The numbers with minus (-) and plus (+) signs indicate the positions of the upstream sequences and the downstream sequences from the adenines of start codons, respectively. **d**, Carpel number quantifications of WT, *slclv3*, *slclv3 slcle9* mutants compared to the T_1 transgenic plants $gSlCLV3^{SlCLV3}$*-2*, $gSlCLV3^{SlCLE9}$*-2*, $gSlCLE9^{SlCLE9}$*-2*, and $gSlCLE9^{SlCLE9S6G}$. Data are based on at least three independent transgenic lines for each construct. Box plots, 25^{th}-75^{th} percentile; center line, median; whiskers, full data range in **b** and **d**. The letters indicate the significance groups at $P < 0.01$ (One-way ANOVA and Tukey test) in **b** and **d**. Different letters between genotypes indicate significance in **b** and **d** (See Supplementary Data 7 for specific *P* values). Exact sample sizes (n) are shown in **b** and **d**. At least twice experiments were repeated independently with similar results.

Supplementary Tables

Supplementary Table 1. CLE dodecapeptide sequences of SlCLV3 and SlCLE9 homologs

Supplementary Table 2. Differentially expressed genes between petunia WT and *phclv3* from mRNA-seq. For the statistical test, "Wald test" was performed, and adjustments were made for multiple comparison. Significant differential expression was identified using *padj* ≤ 0.05 cut-offs and |log2_ratio| ≥ 1 (See Methods).

Supplementary Table 3. Differentially expressed genes between groundcherry WT and *pgclv3* from mRNA-seq. For the statistical test, "Wald test" was performed, and adjustments were made for multiple comparison. Significant differential expression was identified using *q*-value ≤ 0.01 cut-offs (See Methods).

Supplementary Table 4. Primers used in this study.

Supplementary Data

Supplementary Data 1. CRISPR-generated mutations in this study

Supplementary Data 2. Quantification data for organ numbers in this study.

Supplementary Data 3. Quantification data for meristem size from Fig. 1, 2 and 3.

Supplementary Data 4. Normalized counts from mRNA-seq for Fig.2 and 3.

Supplementary Data 5. Syntenic region of *SlCLV3* homologs, defined by Conservatory orthogroups.

Supplementary Data 6. Syntenic region of *SlCLE9* homologs, defined by Conservatory orthogroups.

Supplementary Data 7. Exact *P*-values in this study (one-way ANOVA with Tukey test).

Supplementary Data 8. Sequencing trace files.

Supplementary Data 9. Petunia *PhCLE9* sequence.

Supplementary Data 10. CLV1 homolog sequences.

Figure 1

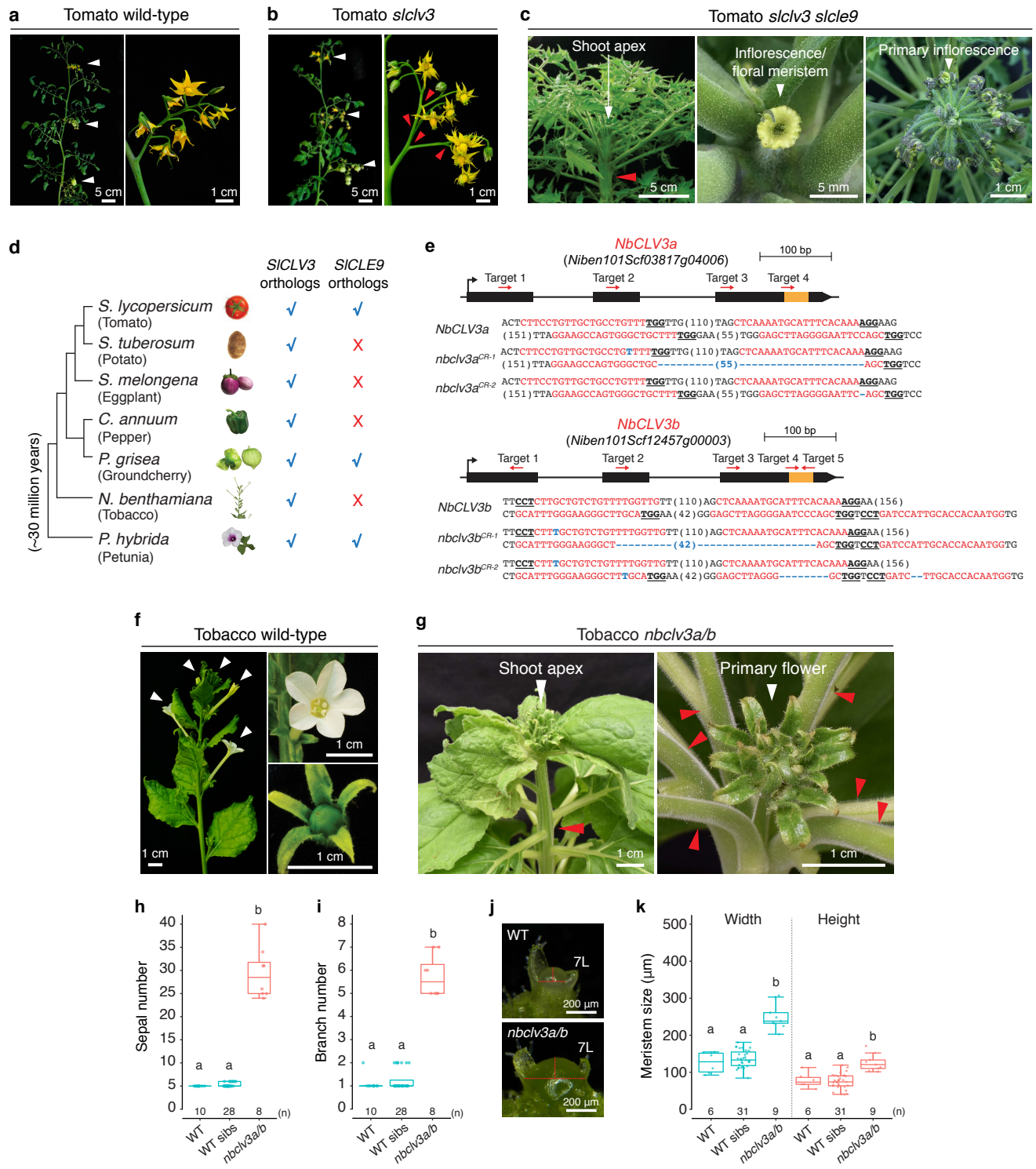

a
Tomato wild-type
5 cm
1 cm
b
Tomato slclv3
5 cm
1 cm
c
Tomato slclv3 slcle9
Shoot apex
5 cm
Inflorescence/
floral meristem
5 mm
Primary inflorescence
1 cm
d
SlCLV3 orthologs
SlCLE9 orthologs
(~30 million years)
S. lycopersicum (Tomato) √ √
S. tuberosum (Potato) √ X
S. melongena (Eggplant) √ X
C. annuum (Pepper) √ X
P. grisea (Groundcherry) √ √
N. benthamiana (Tobacco) √ X
P. hybrida (Petunia) √ √
e
NbCLV3a
(Niben101Scf03817g04006)
100 bp
Target 1
Target 2
Target 3
Target 4
NbCLV3a
nbclv3a^CR-1
nbclv3a^CR-2
NbCLV3b
(Niben101Scf12457g00003)
100 bp
Target 1
Target 2
Target 3
Target 4
Target 5
NbCLV3b
nbclv3b^CR-1
nbclv3b^CR-2
f
Tobacco wild-type
1 cm
1 cm
1 cm
g
Tobacco nbclv3a/b
Shoot apex
1 cm
Primary flower
1 cm
h
Sepal number
a
a
b
10
28
8
(n)
WT
WT sibs
nbclv3a/b
i
Branch number
a
a
b
10
28
8
(n)
WT
WT sibs
nbclv3a/b
j
WT
7L
200 μm
nbclv3a/b
7L
200 μm
k
Width
Height
Meristem size (μm)
a
a
b
a
a
b
6
31
9
6
31
9
(n)
WT
WT sibs
nbclv3a/b
WT
WT sibs
nbclv3a/b

Figure 2

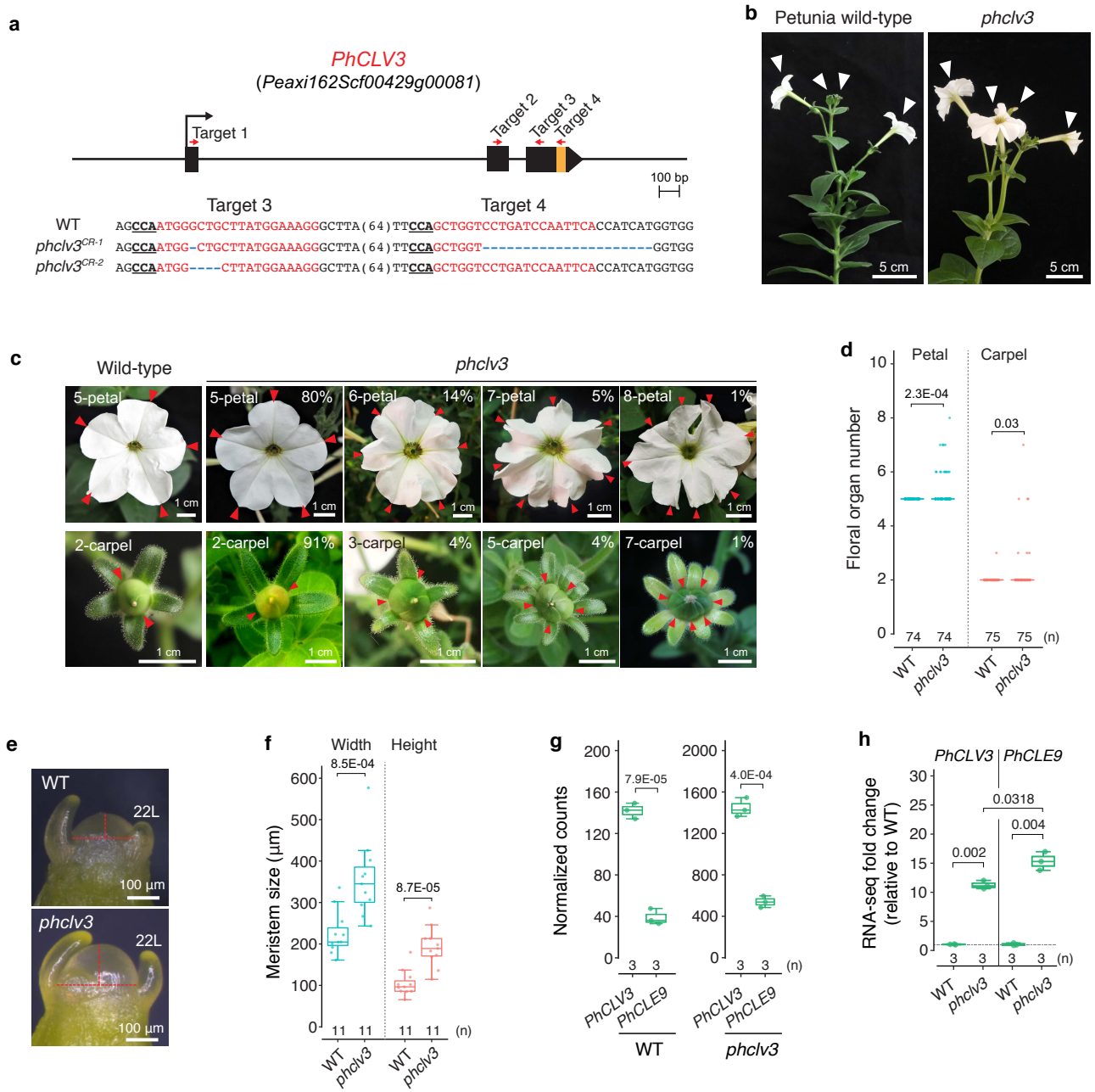

a
PhCLV3
(Peaxi162Scf00429g00081)
Target 1
Target 2
Target 3
Target 4
100 bp
Target 3
Target 4
WT AGCCAATGGGCTGCTTATGGAAAGGGCTTA(64)TTCCAGCTGGTCCTGATCCAATTCACCATCATGGTGG
phclv3CR-1 AGCCAATGG-CTGCTTATGGAAAGGGCTTA(64)TTCCAGCTGGT----------GGTGG
phclv3CR-2 AGCCAATGG----CTTATGGAAAGGGCTTA(64)TTCCAGCTGGTCCTGATCCAATTCACCATCATGGTGG
b
Petunia wild-type
phclv3
5 cm
5 cm
c
Wild-type
phclv3
5-petal
5-petal 80%
6-petal 14%
7-petal 5%
8-petal 1%
2-carpel
2-carpel 91%
3-carpel 4%
5-carpel 4%
7-carpel 1%
1 cm
d
Petal
Carpel
Floral organ number
2.3E-04
0.03
74 74 75 75 (n)
WT phclv3 WT phclv3
e
WT
22L
100 µm
phclv3
22L
100 µm
f
Width
Height
Meristem size (µm)
8.5E-04
8.7E-05
11 11 11 11 (n)
WT phclv3 WT phclv3
g
Normalized counts
7.9E-05
4.0E-04
3 3 3 3 (n)
PhCLV3 PhCLE9 PhCLV3 PhCLE9
WT
phclv3
h
PhCLV3
PhCLE9
RNA-seq fold change (relative to WT)
0.0318
0.004
0.002
3 3 3 3 (n)
WT phclv3 WT phclv3

Figure 3

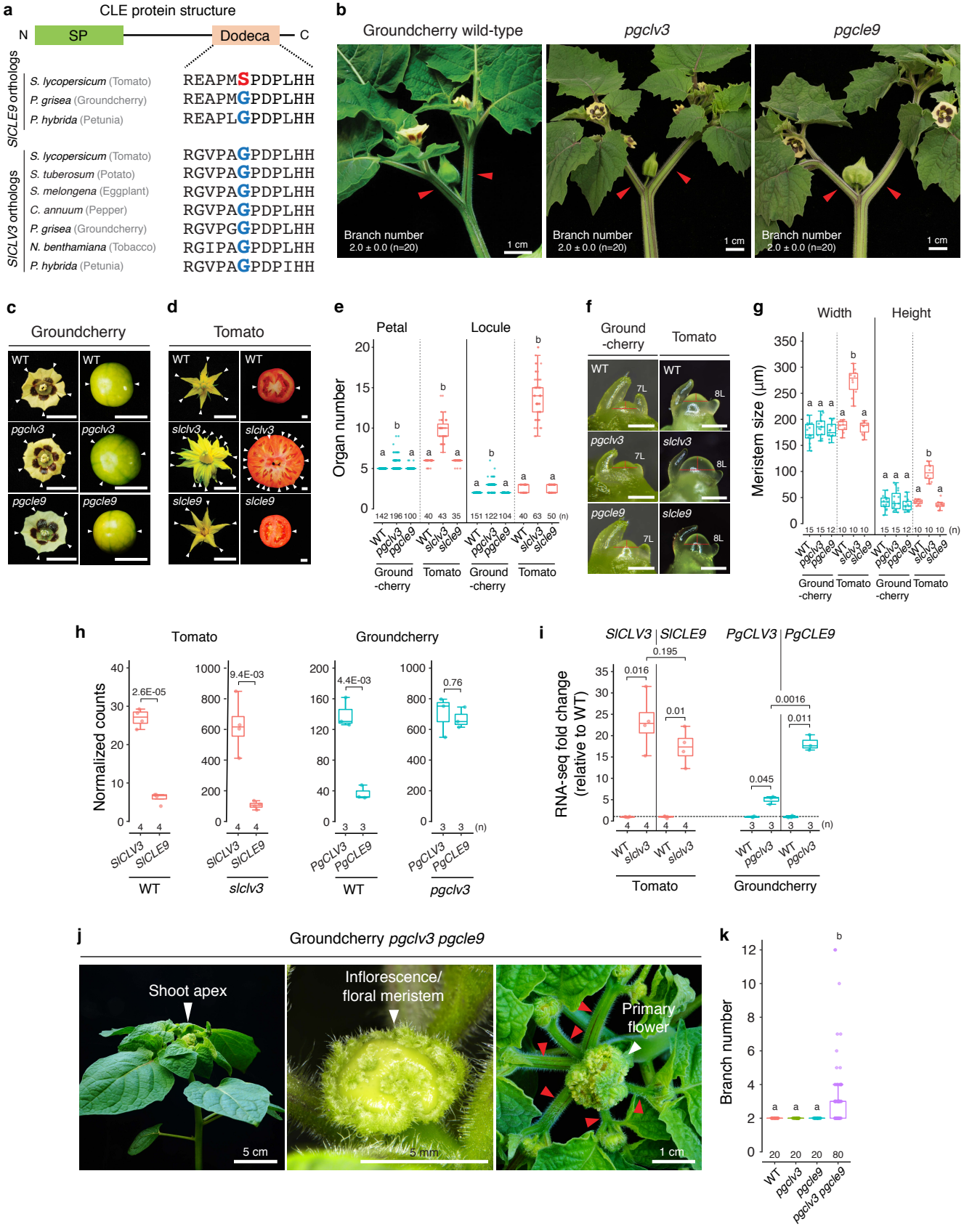
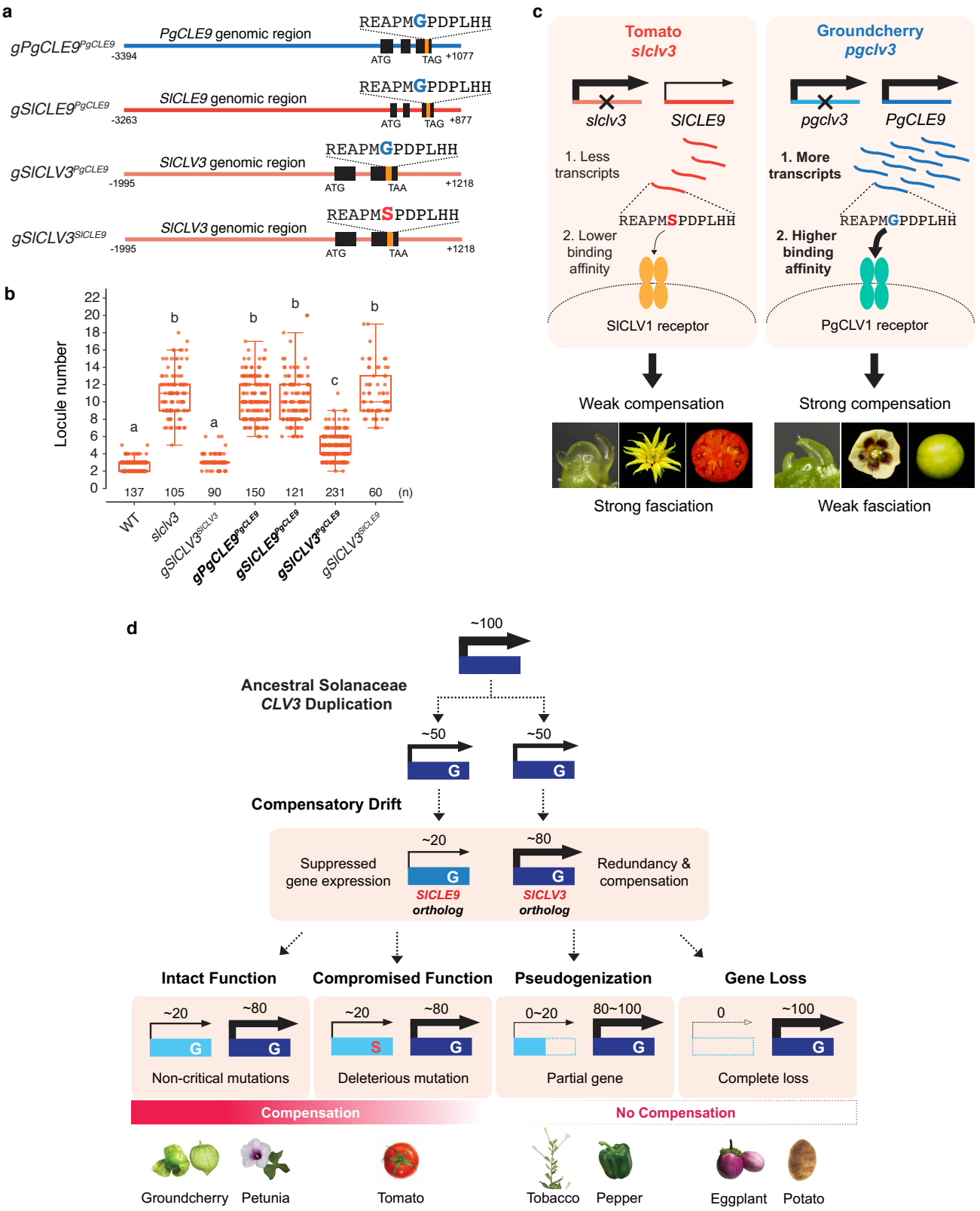

a
CLE protein structure
N
SP
Dodeca
C
SlCLE9 orthologs
S. lycopersicum (Tomato) REAPMSPDPLHH
P. grisea (Groundcherry) REAPMGPDPLHH
P. hybrida (Petunia) REAPLGPDPLHH
SlCLV3 orthologs
S. lycopersicum (Tomato) RGVPAGPDPLHH
S. tuberosum (Potato) RGVPAGPDPLHH
S. melongena (Eggplant) RGVPAGPDPLHH
C. annuum (Pepper) RGVPAGPDPLHH
P. grisea (Groundcherry) RGVPGGPDPLHH
N. benthamiana (Tobacco) RGIPAGPDPLHH
P. hybrida (Petunia) RGVPAGPDPIHH
b
Groundcherry wild-type
pgclv3
pgcle9
Branch number
2.0 ± 0.0 (n=20)
1 cm
c
Groundcherry
d
Tomato
WT
pgclv3
pgcle9
slclv3
slcle9
e
Petal
Locule
Organ number
Ground-cherry
Tomato
f
7L
8L
g
Width
Height
Meristem size (µm)
h
Tomato
Groundcherry
Normalized counts
2.6E-05
9.4E-03
4.4E-03
0.76
SlCLV3
SlCLE9
PgCLV3
PgCLE9
i
SlCLV3 | SlCLE9
PgCLV3 | PgCLE9
RNA-seq fold change (relative to WT)
0.016
0.195
0.01
0.0016
0.011
0.045
j
Groundcherry pgclv3 pgcle9
Shoot apex
Inflorescence/ floral meristem
Primary flower
5 cm
5 mm
1 cm
k
Branch number
pgclv3 pgcle9

Figure 4

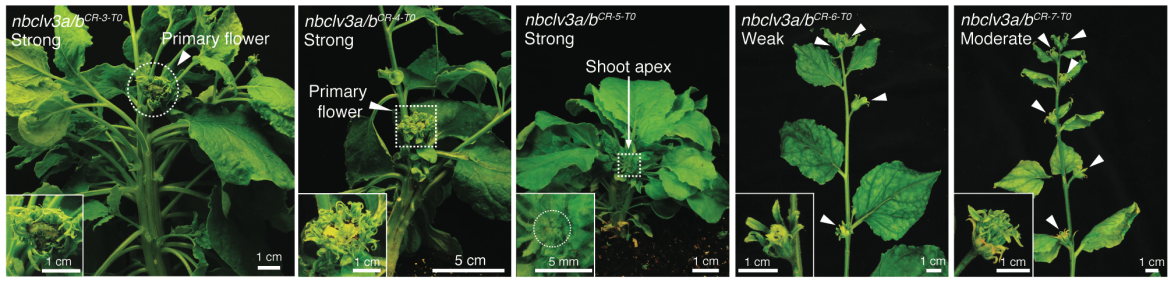

a
gPgCLE9PgCLE9
PgCLE9 genomic region
REAPMGPDPLHH
-3394
ATG
TAG
+1077
gSlCLE9PgCLE9
SlCLE9 genomic region
REAPMGPDPLHH
-3263
ATG
TAG
+877
gSlCLV3PgCLE9
SlCLV3 genomic region
REAPMGPDPLHH
-1995
ATG
TAA
+1218
gSlCLV3SlCLE9
SlCLV3 genomic region
REAPMSPDPLHH
-1995
ATG
TAA
+1218
b
Locule number
22
20
18
16
14
12
10
8
6
4
2
0
a
b
a
b
b
c
b
137
105
90
150
121
231
60
(n)
WT
slclv3
gSlCLV3SlCLV3
gPgCLE9PgCLE9
gSlCLE9PgCLE9
gSlCLV3PgCLE9
gSlCLV3SlCLE9
c
Tomato
slclv3
slclv3
SlCLE9
1. Less transcripts
REAPMSPDPLHH
2. Lower binding affinity
SlCLV1 receptor
Weak compensation
Strong fasciation
Groundcherry
pgclv3
pgclv3
PgCLE9
1. More transcripts
REAPMGPDPLHH
2. Higher binding affinity
PgCLV1 receptor
Strong compensation
Weak fasciation
d
~100
Ancestral Solanaceae
CLV3 Duplication
~50
~50
G
G
Compensatory Drift
Suppressed gene expression
~20
~80
G
G
SlCLE9 ortholog
SlCLV3 ortholog
Redundancy & compensation
Intact Function
Compromised Function
Pseudogenization
Gene Loss
~20
~80
G
G
Non-critical mutations
~20
~80
S
G
Deleterious mutation
0~20
80~100
G
Partial gene
0
~100
G
Complete loss
Compensation
No Compensation
Groundcherry
Petunia
Tomato
Tobacco
Pepper
Eggplant
Potato

Extended Data Fig. 1

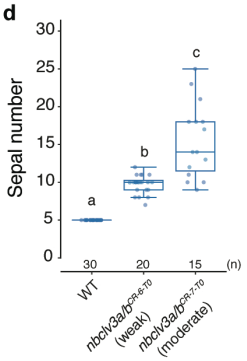
a

b

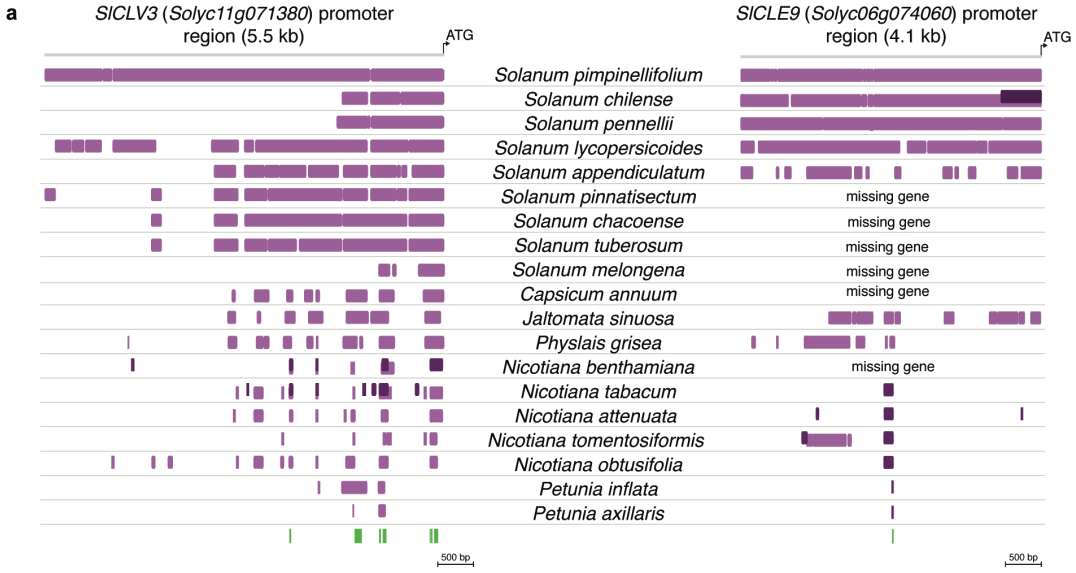
c

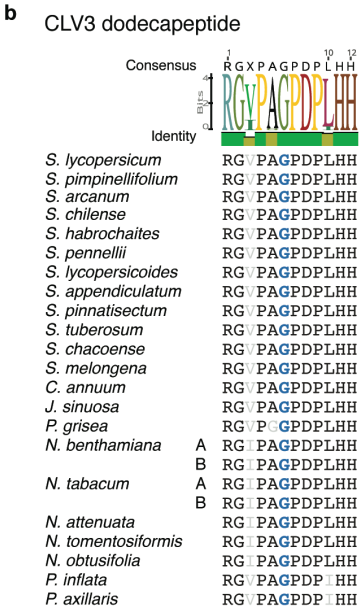


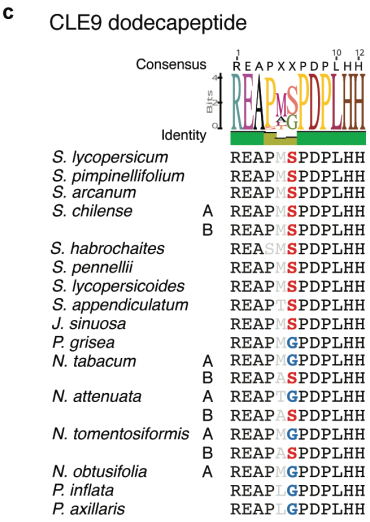
d



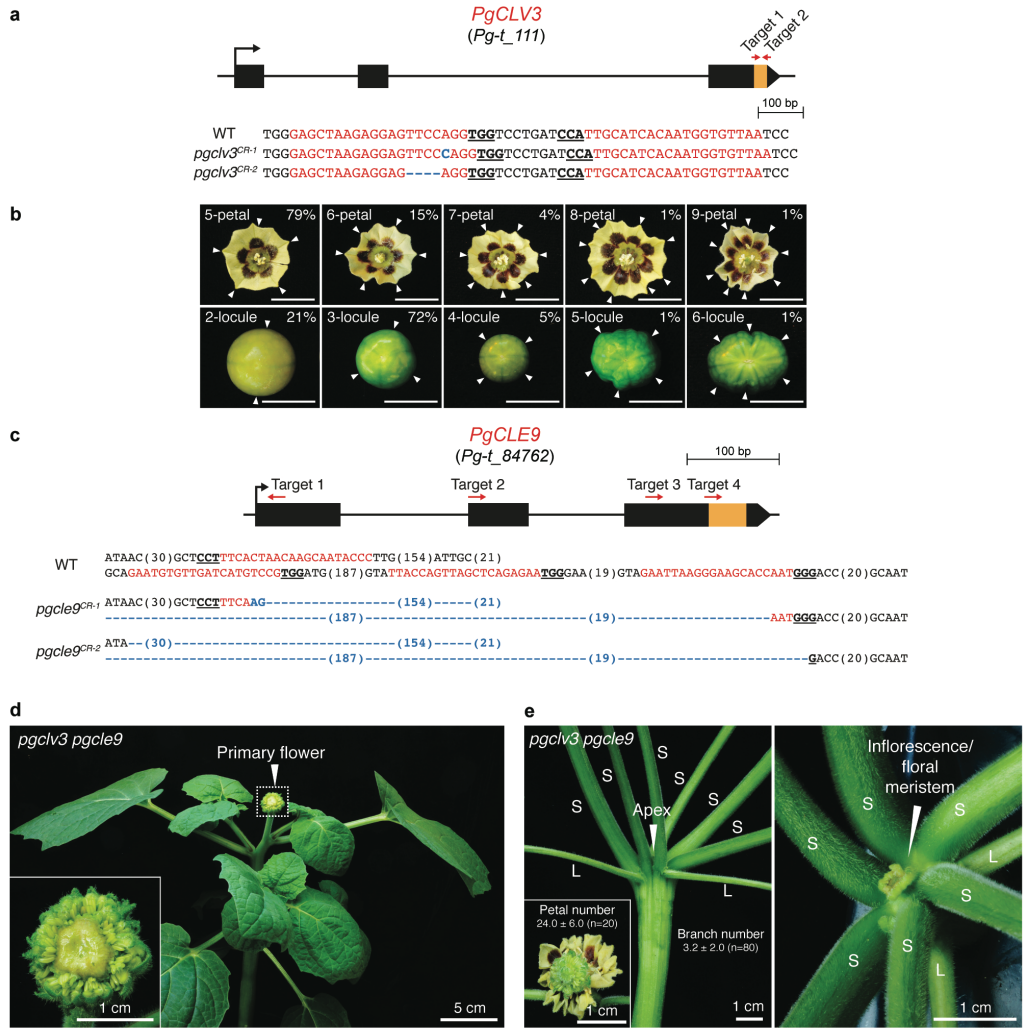
Extended Data Fig. 2







Extended Data Fig. 3



Extended Data Fig. 4

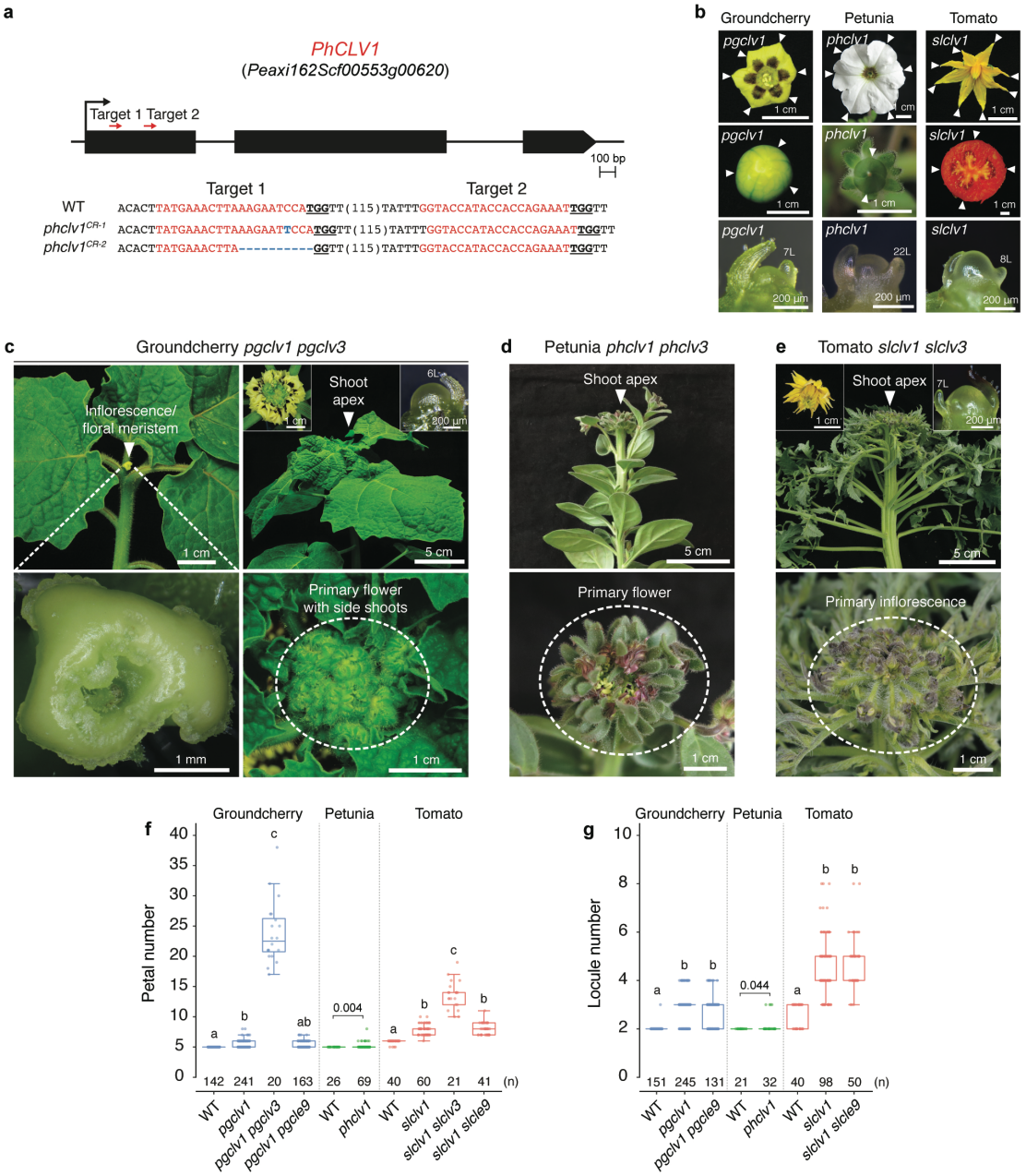
a

```
SlCLV1          PSEFGNISTLKLLDLGNCNLDGEVPPSLGNLKKLHSLFLQVNRLTGHIPSELSGLESLMS 293
StCLV1          PSEFGNISTLKLLDLGNCNLDGEVPPSLGNLKKLHTLFLQVNRLTGRIPSELSGLESLMS 292
SmCLV1          PSEFGSISTLKLLDLGNCNLEGEIPPSLGNLKKLHTLFLQMNRLTGHIPTELSGLESLMS 292
CaCLV1          PPEFGNISTLKLLDLGSCNLDGEIPPSLANLKKLHSLFLQMNRLTGRIPSELSGLYSLMS 295
PgCLV1          PPEFGSISTLKYLDLGSCNLDGEIPPSLGNLKKMHTLFLQVNRLTGRIPSELSGLESLMS 293
NbCLV1          PTEFGSISTLKLLDLANCNLDGEIPPSLGNLKKLHSLFLHANRLTGHIPSELSGLESLMS 296
PhCLV1          PPEFGSITTLKLLDLGSCNLDGEIPASLGNLKKLHSLFLQMNRLTGYIPPELSGLESLMS 290
                * ***. :*** ***..***:**:* **.****:*:***: ***** ** ***** ****
```

b

```
                                                D                F
Solyc04g081590.2 (SlCLV1)       EELRLGYYNSYEGGIPSE-FGNISTLKLLDLGNCNLDGEVPPSLGNLKKLHSLFLQVNRL   277
AT1G75820                       REMYIGYYNSYTGGVPPE-FGGLTKLEILDMASCTLTGEIPTSLSNLKHLHTLFLHINNL   277
GSMUA_Achr3G26680_001           ----------YEGGIPWE-FGRLSSLVRLDMAGCRLSGTLPASLGQLKRLDSLFLQINRL   249
Bradi1g30160                    EDLYLGYFNQYDGGVPPE-FGELASLVRLDMSSCNLTGPVPPELGKLSKLQTLFLLWNRL   339
Brast07G235600                  QDLYLGYYNQYDGGVPPE-FGALGSLIRLDMSSCNLTGPIPPELGMLSNLETLFLQWNRL   288
LOC_Os06g50340                  REMYIGYYNQYDGGVPPE-FGDLGALLRLDMSSCNLTGPVPPELGRLQRLDTLFLQWNRL   281
Sevir.4G294000                  REMYIGYYNQYTGGVPPE-FGDLRSLVRLDISSCNLTGPVPPELARLTQLDTLFLSINQL   318
Seita.4G281800                  REMYIGYYNQYTGGVPPE-FGDLRSLVRLDISSCNLTGPVPPELARLTQLDTLFLSINQL   318
Pahal.D00165                    REMYIGYFNQYTGGVPPE-FGDLRSLVRLDMSSCNLTGPVPLELARLTQLDTLFLSINQL   315
Pavir.Db00153                   REMYIGYFNQYTGGVPPE-FGDLRSLVRLDMSSCNLTGPVPPELARLTQLDTLFLSINQL   318
Sobic.010G267700                REMYVGYYNQYSGGVPPE-FGDLQSLVRLDMSSCTLTGPIPPELARLSRLDTLFLSMNQL   289
GRMZM2G300133                   REMYVGYYNQYSGGVPRE-FGALQSLVRLDMSSCTLTGPIPPELARLSRLDTLFLALNQL   288
MDP0000280399                   KELYLGYYNNYDGGIPPE-FGSLPLLKVLDMSSCNLIGKIPTSLSLLKNLHSLFLQVNRL   674
AHYPO_018678                    NMLFLGYYNTFSGGIPSE-FGSLSSLKLLDMASCNLSGEIPKTLGNLKNVHTLFLQRNQL   279
Migut.N03171                    QELYLGYFNTYDGGIPPA-FGSISTLQLLDLAMCNLTGEIPASLGNLKHLHSLFLQVNNL   278
Migut.C00856                    LELYLGYYNTYSGGIPPE-FGSISSLQLLDLGMCNLTGEIPATLGNLKHLHTLFLQVNNL   258
DCAR_022991                     QILRLGYYNMYLGGIPSE-LGTLSDLRLLDLGGCNLTGEIPASLGNLKLLHTLFLQYNNL   255
PGSC0003DMG400009941            EELRLGYYNSYEGGIPSE-FGNISTLKLLDLGNCNLDGEVPPSLGNLKKLHTLFLQVNRL   276
Eucgr.H00964                    RGLYLGYFNAFDGGIPAE-FGSLKELRILDMASCGLSGEIPASLGELKKLDSLFLHLNKL   282
Eucgr.H00963                    QWLYLGYFNTYDGEIPAE-FGSMKELRRLDLASCGLSGEIPVSLSELKKLDSLFLQWNNL   269
Phvul.011G042000                KYLKLGYNNAYEGGIPPE-FGAMKSLIYLDLSSCNLSGEIPPSLSSLKKLDTLFLQMNNL   296
Glyma.11G114100                 RILKLGYNNAYEGGIPPE-FGTMESLKYLDLSSCNLSGEIPPSLANMRNLDTLFLQMNNL   279
Glyma.12G040000                 RYLKLGYNNAYEGGIPPE-FGSMKSLRYLDLSSCNLSGEIPPSLANLTNLDTLFLQINNL   279
Aqcoe6G222600                   QQLYLGYYNAYEGGIPPE-FGSFESLRLLDLGSCNLSGEIPASLGGLKLLDTLFLQFNHL   275
Prupe.6G163000                  KELYVGYFNSFDGGIPPE-LGSLTWLQVLDLASCNLSGSIPRSLGLLKHLRSLFLQVNCL   280
Kaladp0068s0368                 EQMYVGYFNVYSSGIPPE-FGSIISLRILDMASCNLSGEIPATLGKLKNLDTLFLQVNNF   278
Kalax.0183s0036                 EQMYVGYFNVYSSGIPPE-FGSITSLRILDMASCNLSGEIPATLGKLKNLDTLFLQVNNF   278
Bol027692                       KEMYVGYFNSYTGGVPPE-FGELTNLEVLDMASCTLTGEIPTTLSNLKHLHTLFLHINNL   284
Brara.G03381                    KEMYVGYFNSYTGGVPPE-FGELSNLEVLDMASCTLTGEIPTTLSNLKHLHTLFLHINNL   284
Araha.9358s0001                 KEMYIGYYNSYTGGVPPE-FGGLTKLEILDMASCTLTGEIPTSLSNLKHLHTLFLHINNL   277
AL2G35810                       KEMYIGYYNSYTGGIPPE-FGGLTKLEILDMASCTLTGEIPTSLSNLKHLHTLFLHVNNL   277
Thhalv10018069m.g               REMYVGYYNSYTGGVPPE-FGGLTKLEVLDMASCTLTGEIPTTLSNLKHLHTLFLHINNL   286
Bostr.20129s0016                REMYVGYYNSYTGGVPPE-FGGLTKLEILDMASCTLTGEIPTSLSNLKHLNTLFLHINNL   285
Cagra.0799s0053                 REMYVGYFNSYTGGVPPE-FGGLTKLQILDMASCTLTGEIPTSLSNLKHLHTLFLHINNL   288
Carubv10019714m.g               REMYVGYFNSYTGGVPPE-FGGLTKLQILDMASCTLTGEIPTSLSNLKHLHTLFLHINNL   287
MDP0000804929                   RELYVGYYNSYSGGIPPE-LGSLSSLQILDMGSCNLVGPIPTTLSLLKHLHTLFLQVNRL   275
gene08548-v1.0-hybrid           KEMYVGYFNSYDGGIPPE-LGSLSSLRVLDMASCNLTGTIPISLSNLKHLHSLFLQINQL   274
Prupe.1G363300                  KEMYVGYFNSYDGGIPPE-LGSLSSLQVLDMASCNLSGTIPTNLSLLKNLNSLFLQVNRL   278
GSVIVG01009941001               QGLFLGYFNIYEGGIPPE-LGLLSSLRVLDLGSCNLTGEIPPSLGRLKMLHSLFLQLNQL   276
Potri.002G019900                KSLCIGYYNHYEGGIPPE-FGSLSNLELLDMGSCNLNGEIPSTLGQLTHLHSLFLQFNNL   274
SapurV1A.0025s0150              KSLSIGYFNHYEGGIPPE-FGSLSSLELLDMGSCNLNGEIPSTLGQLTRLQSLFLQFNNL   274
Potri.005G241500                KSLCVGYFNRYEGSIPPE-FGSLSNLELLDMASCNLDGEIPSALSQLTHLHSLFLQVNNL   273
SapurV1A.0384s0100              KSLCVGYFNHYEGIIPPE-FGSLVNLELLDMASCNLNGEIPATLGQLTRLHSLFLQVNNL   274
Manes.05G145600                 RSLYLGYYSSYEGGIPPE-FGFLSSLEVLDMAFCNLTGEIPSTLGLLKRLHTLFLQMNNL   275
30170.t000788                   RKLYLGYFNSWEGGIPPE-FGSLSSLEILDMAQSNLSGEIPPSLGQLKNLNSLFLQMNRL   278
evm.TU.supercontig_26.309       KALFIGYSNLYNGGVPRE-FGDLSELQILDMSSCNITGEIPTSLSNLKHLHSLFLQMNNL   276
Ciclev10000156m.g               REMYIGYFNTYTGGISPE-FGALTQLQVLDMASCNISGEIPTSLSRLKLLHSLFLQMNKL   277
orange1.1g002010m.g             REMYIGYFNTYTGGIPPG-FGALTQLQVLDMASCNISGEIPTSLSRLKLLHSLFLQMNKL   277
Gorai.005G112100                KYLVIGYFNAYDGGIPPE-YGSLSQLELLDMASCNITGEIPSSLSNLKHLHSLFLQLNRL   277
Thecc1EG034252                  KEMYIGYFNAYVGEIPPE-FGTLSQLQVLDMASCNLTGEIPVSLSNLKHLHTLFLQLNRL   277
Medtr4g070970                   KELQLGYENAYSGGIPPE-LGSIKSLRYLEISNANLTGEIPPSLGNLENLDSLFLQMNNL   241
Tp57577_TGAC_v2_gene30515       KELRLGYNNAYEGGVPPE-FGSMKYLRYLEMPSCNLTGEIPPSLGNLENLDSLFLQGNNL   275
Lus10040592.g                   KELYLGYFNSFSGGIPSGMFQGLTSLRVLDMASCNLSGEIPPSLGQLKNLRACYLQLNHF   281
```

Extended Data Fig. 5



Extended Data Fig. 6